12-1989

"What Injures Royalty" Surnames, Inbreeding, and Genetic Structure in Sevier County, Tennessee: 1856-1905

Joseph C. Lewelling

To the Graduate Council:

I am submitting herewith a thesis written by Joseph C. Lewelling entitled ""What Injures Royalty" Surnames, Inbreeding, and Genetic Structure in Sevier County, Tennessee: 1856-1905." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Arts, with a major in Anthropology.

Richard L. Jantz, Major Professor

We have read this thesis and recommend its acceptance:

P. Willey, Benita Howell

Accepted for the Council:
Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting herewith a thesis written by Joseph C. Lewelling entitled "'What Injures Royalty', Surnames, Inbreeding, and Genetic Structure in Sevier County, Tennessee: 1856-1905." I have examined the final copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Arts, with a major in Anthropology.

Richard L. Jantz, Major Professor

We have read this thesis
and recommend its acceptance:

P S Willey

Benita J. Howell

Accepted for the Council:

Vice Provost
and Dean of The Graduate School

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Master's degree at The University of Tennessee, Knoxville, I agree that the Library shall make it available to borrowers under rules of the Library. Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgement of the source is made.

Permission for extensive quotation from or reproduction of this thesis may be granted by my major professor, or in his absence, by the Head of Interlibrary Services, when, in the opinion of either, the proposed use of the material is for scholarly purposes. Any copying or use of the material in this thesis for financial gain shall not be allowed without my written permission.

Signature Joseph C. Lewelling

Date November 13, 1989

"WHAT INJURES ROYALTY"

SURNAMES, INBREEDING, AND GENETIC STRUCTURE

IN SEVIER COUNTY, TENNESSEE: 1856-1905

A Thesis

Presented for the

Master of Arts

Degree

The University of Tennessee, Knoxville

Joseph C. Lewelling

December 1989

For my father, Carl Lewelling, Jr.

ACKNOWLEDGMENTS

Many individuals have contributed to this project. In particular I am grateful to my major professor, Richard Jantz, and to Benita Howell and P. Willey, the members of my committee. Their guidance, consultation, and encouragement have been essential to the completion of my thesis, and are greatly appreciated. Also, I am indebted to Sevier County's historian, Beulah Linn, whose knowledge and assistance were invaluable to my research. Finally, I want to express my appreciation to my wife, Vickie, for her considerable patience during the extended research and writing of this thesis.

Any errors remaining in this work are solely my own.

ABSTRACT

Inbreeding in Southern Appalachia has been a topic of discourse for both local color writers and academicians since the late 19th century, but only a few researchers have aspired to measure inbreeding or describe the genetic structure of Appalachian populations. This study attempts to assess the genetic structure in one small section of Appalachia--Sevier County, Tennessee.

Because surnames, like alleles, are inherited from a parent, their distribution in a population can give clues about the genetic structure of that population--including the degree of inbreeding. Census and marriage records from 1856-1905 were examined using several methods of surname analysis, and estimates of kinship and inbreeding were calculated for the county, and for each of the fifteen civil districts within the county. Genetic distances were calculated between all districts, and repeated pairs analysis was applied to the marriage records.

Inbreeding and kinship coefficients in Sevier County were lower than those observed in true isolate populations, but the large nonrandom components of the inbreeding coefficients and in the repeated pairs analysis indicate subdivision of the breeding population. Kinship was notably higher in the mountain districts than in lower-lying districts indicating some geographic constraint on gene flow. Cultural and economic influences on gene flow were also examined.

TABLE OF CONTENTS

LIST OF TABLES

TABLE		PAGE

LIST OF FIGURES

CHAPTER 1

Inbreeding and the Southern Highlander

The people who dwelled among the rugged mountains of eastern Tennessee were often viewed by outsiders as different, a class of people set apart from their American contemporaries by distinct speech, customs, and living conditions. To some they were Anglo-America's Noble Savages--living artifacts of a Scotch-Irish heritage, who existed in a quiet and simple balance with nature (Caudill 1962:77; Dunn 1988:xiii). Others had a less romantic view, believing the mountaineers to be a degraded lot, uncivilized and basically uncivil (Dunn 1983:xiii; Hirsch 1928).

To the outsiders holding this latter opinion, the degradation grew out of isolation and poverty, and was clearly discernable in the mountaineers' cruelty, in their reclusive nature, and especially, in their ignorance. This portrait, originally drawn in the late 19th and early 20th centuries by local-color writers, also found support among certain academicians.

> The educational facilities of these people were limited and none of them have probably passed through a high school, nor is their work such as would require any great mental activity. The head in this group shows a decidedly lower average in size...(Hrdlicka 1925:184)

Although social conditions were initially thought to be responsible for the Highlanders' supposed low intelligence, the popularity of eugenics in the early years of this century gave rise to a new theory. After a degree of clannishness and frequent consanguineous marriages were noted among the mountaineers, it was postulated that their intellect was hindered by genetic inadequacy--the result of generations of seclusion and intermarriage. As one early observer put it:

> In some of the more isolated mountain districts, there has been, owing to their isolation, too much intermarriage even; and what injures European royalty does not improve mountain society (Wilson 1914:67-68).

Intermarriage quickly became a part of the color writers' depictions (Kephart 1976:297), and scientists hurriedly began examining these "inbred" populations. Their studies revealed a variety of supposed "mountain traits," ranging from decreased variability (Carter 1928) to mental retardation and congenital deformation (Manne 1936; Hirsch 1928). The appearance of such traits was seen as "the evil consequences of inbreeding of persons closely akin..."(Kephart 1913:297).

But few of these early studies attempted to measure inbreeding in any precise way, or to compare it with levels found in other groups. A researcher who undertook such a comparison discovered an elevated rate of intermarriage in one mountain community, but a rate far less than that occurring among forty European royal families (Carter 1928).

Still, the notion that the Southern Highlanders were "fantastically inbred" (Caudill 1962:84) persisted. The same genetic consequences continued to be asserted, from mental deficiency (Weyl 1974) to a homogeneity that reached beyond the mere physical:

> ...at the end of three or four generations the people tended to look alike, talk with the similar inflections, and share the same preferences in food, fiddle tunes, and politics (Caudill 1962:12).

Recently the validity of this impression has been questioned. Researchers have claimed that the rate of intermarriage varies from community to community within Appalachia, and that inbreeding in these mountains is not significantly more prevalent than inbreeding in other regions (Tincher 1980). Additionally, it has been maintained that the original gene pool was both large and "good" so that inbreeding at a gross level would have had little damaging effect (Trout 1984:23).

Even low rates of inbreeding, however, may have profound consequences for individuals. The homozygotic occurrence of particular alleles has been linked with a number of congenital abnormalities and conditions. Consanguineous matings result in decreased heterozygosity, so a genetic disease resulting from homozygosity of an allele might become more prevalent in inbred populations. This situation may worsen if genetic drift results in a random increase in the deleterious alleles.

Within extremely inbred lineages, an inadequately understood phenomenon, inbreeding depression, may also occur. This "depression" is well known in commercial plant and animal production where inbred parent stocks consist of "sorry looking specimens", while the hybrids of separated strains possess "superior qualities" (Brues 1977:70).

A possible explanation of this "hybrid vigor" is that heterozygotes may sometimes enjoy an advantage over homozygotes because the different alleles operate synergistically rather than independently or exclusively. A recessive allele that does not interfere with the function of a dominant one "may add something by way of versatility or efficiency...without taking anything away" (Brues 1977:70).

Heterosis, the technical name for this phenomenon, is best known in agricultural genetics (Underwood 1979:88), but its consequences in human populations are less certain. Some studies have found a positive correlation between heterozygosity and stature in humans (Hulse 1958; Schreider 1969), but evidence of other such effects is scant.

One result of heterosis should be an increase in, or the maintenance of, heterozygosity within a population. If heterozygosity is advantageous, then selection should work to weed out homozygotes while retaining both alleles in the superior surviving heterozygotes. Inbreeding, as a loss of hetrozygosity, would thus be minimized.

Population inbreeding becomes even more complex if it is defined, not as simply the intermarriage of "closely" related individuals, but as an increase in homozygosity due to the inheritance of identical genes from common ancestors. By this definition, it is a function of several factors, the most important being population size.

Generally, inbreeding accumulates faster in a population with fewer breeding individuals. An example can be found on Pitcairn Island in the South Pacific, which was settled in 1790 by Fletcher Christian, the mutineers of the *HMS Bounty*, and their Tahitian mates. Two centuries of isolation and intermarriage produced a descendant population that is greatly inbred (Lasker 1985:28-29).

If a founding population consisted of related individuals, then homozygosity among descendants will be even greater. The Hutterites studied by Crow and Mange (1965) were already closely related upon their arrival in America in 1879. Less than half of their estimated inbreeding accumulated after that time; the remainder can be attributed to intermarriage before their arrival.

Because the number of potential breeders is finite, while the number of potential ancestors approaches infinity, some distant intermarriage occurs even in large populations. If a population has a long history of little or no immigration, then even distant intermarriage can, over an extended number of generations, result in a loss of heterozygosity. By this process even large breeding groups may become highly inbred.

Homozygosity in a large population may also increase if that population is subdivided into smaller groups or endogamous clusters. In such situations genetic barriers exist between the separate communities or lineages that make up the larger population, effectively partitioning it into smaller breeding isolates. Homozygosity within these isolates then increases because alleles are not randomly distributed throughout the population.

While breeding subdivision does increase individual homozygosity, it does not necessarily reduce overall genetic variability. Different endogamous clusters may become homozygous for different alleles, increasing between-cluster variation, while decreasing variation within the smaller breeding units. In fact, the homogenizing effects of genetic drift may be lessened if alleles randomly lost to the overall gene pool are retained in a subpopulation, and then occasionally reintroduced into the larger population.

To gauge the possible impact of inbreeding on a particular population, it is necessary to have at least relative measures of the genetic and demographic structure of that population. For the mountains of Tennessee and the region of Southern Appalachia, however, this structure is not well known. Population size, isolation, and subdivision have been subjects of speculation, but only a few studies of

Appalachian peoples have furnished reliable information on genetic structure (Tincher 1980; Kirkland and Jantz 1977; Politzer and Brown 1969; Carter 1928). Even with these studies, small sample sizes, limited geographical areas, and disparate results preclude any generalization about the region as a whole.

Additional research is obviously needed. Given the possible variation of inbreeding rates throughout the region (Tincher 1980:28), the research must take the form of multiple studies in different localities. Ideally, historical groups would be investigated since Appalachia has undergone significant demographic change over recent decades. A variety of methods might be employed to estimate both rates and patterns of inbreeding. Also, the influence of various geographical or social phenomena could be examined. Finally, these studies should provide some degree of historical depth in order that changes over time might be measured. Until such research is completed and a sizable body of data is compiled, the inbred Southern Mountaineer will remain largely a figure of fiction and speculation.

This project attempts to address that research need and contribute to that body of data. By closely examining the genetic structure of a small, isolated area, it may be possible to increase the knowledge of Appalachian genetics as a whole. The area chosen for this study is Sevier County, Tennessee, an expanse of land rising from the French Broad

River and climbing the rugged western flank of the Great Smoky Mountains.

CHAPTER 2

Geography and Settlement

The French Broad River flows west from the high mountains along the Tennessee-North Carolina border. As it leaves these mountains, it turns south, toward a confluence with the larger Holston River just east of Knoxville, Tennessee.

In the late 18th century, when most rivers served as highways for early settlers, the French Broad was a barrier, separating the East Tennessee Valley from the unceded Indian lands to the south. Most of the 597 square miles that eventually became Sevier County lay south of this river in those wild, mountainous, Indian lands. Only a small triangular tract lay to the north (Smoky Mountain Historical Society 1986:1-2). Drained chiefly by branches of the Little Pigeon River, a tributary of the French Broad, this land has been described as three more or less distinct geographical zones (Figure 1).

The northwestern third of the county included the wide low floodplains of the French Broad, Boyds Creek, and the confluence of the Little Pigeon. The topography here was level to rolling and the waterways were generally navigable by flatboat (Smoky Mountain Historical Society 1986:2). The hills were low and gently rounded, while the soil was a rich loam (Smoky Mountain Historical Society 1986:2; 1984:V).

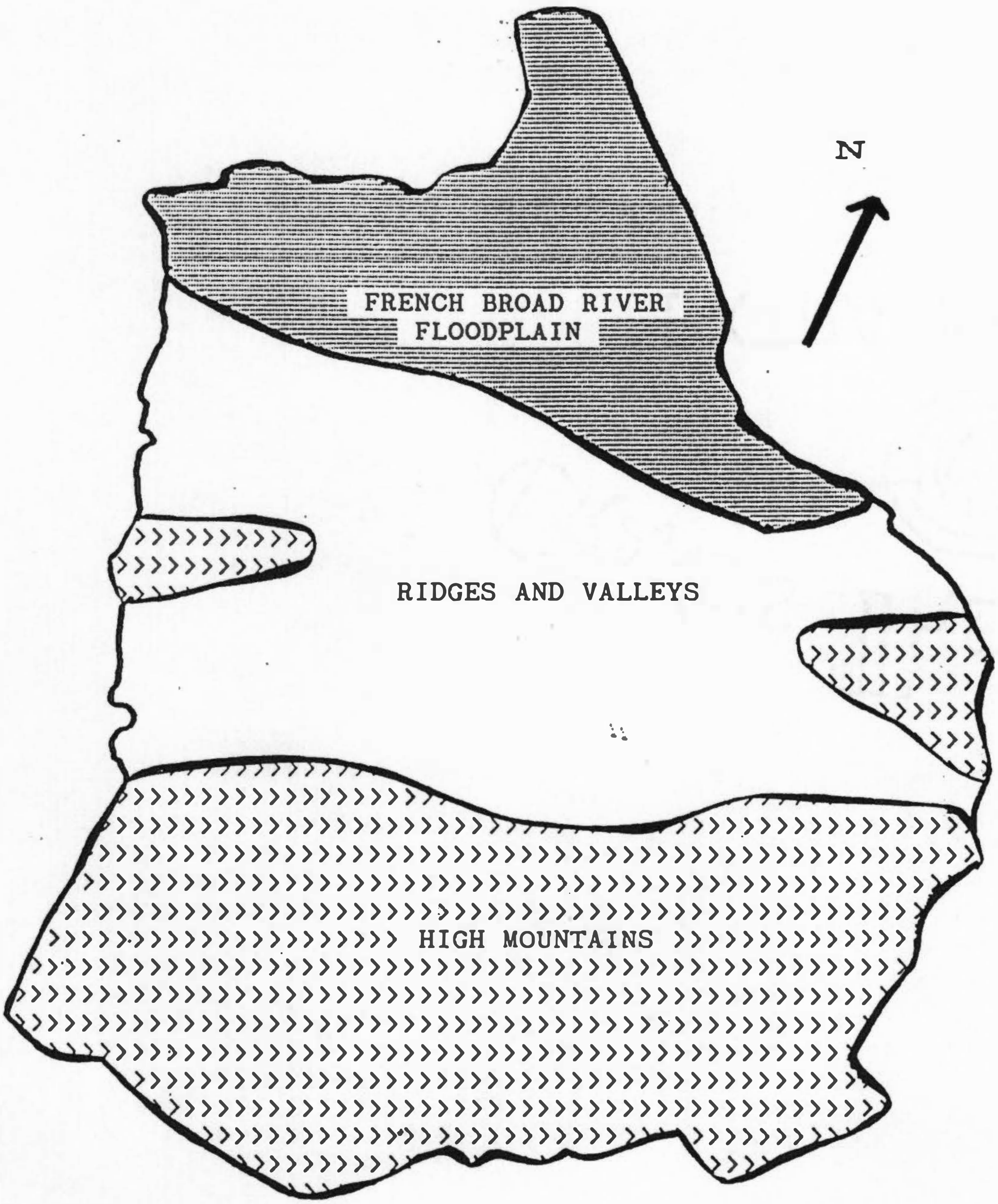


Figure 1. Geographical Zones of Sevier County

The central section of the future county began above the confluence of the Little Pigeon and consisted of the lower drainage basin of that river system. More rugged than the northwestern third, the landscape here was characterized by high hills, ridges, and widely spaced mountains. The valleys narrowed in places, but level tracts were common along the shallow, unnavigable streams. Some land was arable, but in many places the soil was sterile yellow or red clay (Smoky Mountain Historical Society 1984:V; 1986:2).

The landscape of the southeastern third was even more dramatic, with deep narrow gorges, tumbling streams, and high, blue mountains rising above 6000 feet. The soil was generally richer than in the center of the county, though only a few flat acres were scattered among the mountain coves (Smoky Mountain Historical Society 1984:V; 1986:2).

There were, however, no definitive borders for these three sections. The mountains, rivers, and ridges followed irregular courses across the landscape, and each section simply graded into the next. In fact the overall topography of the county was and is best described as a wavering gradient from the low, rich floodplain in the northwest to the towering mountains in the southeast.

Geography was influential, both historically and culturally, in the settlement of the county. The earliest Euro-Americans came to Sevier County by following the Holston River from North Carolina. They settled first on the triangular tract laying north of the French Broad (Smoky Mountain Historical Society 1984:1). After several years of war with the Cherokees who owned the land, the river was crossed and the southern floodplain occupied.

Small communities, like Boyds Creek and Trundles Crossroads, were established. Sevierville, founded at the Forks of the Little Pigeon River, became the county seat when Tennessee attained statehood in the 1790's (Smoky Mountain Historical Society 1984:V).

With the markets at Knoxville accessible by water and with rich arable land plentiful, the settlers in the northwestern section prospered. Large tracts of land were cleared for farming and grazing. Even a small plantation, with scores of slaves, was operating by the early decades of the 19th century. Sevierville endured with its central location, but many communities, especially those north of the river, were tied to markets and towns in other counties (Smoky Mountain Historical Society 1984:1; 1986:2-3; Cummings 1988:25).

As the more commercially valuable lands became populated, new arrivals were forced into the coves and hills of the central section. Small satellite villages appeared to the south and east of Sevierville. Much bottomland was cleared for farming, but the upper ridges and mountains were left mostly wooded and uninhabited (Smoky Mountain Historical Society 1986:2-3).

Finally, settlers homesteaded among the narrow gorges and steep slopes near the North Carolina border. Clearing land in the valleys, they established small farming communities like Gatlinburg, Greenbriar, and Emerts Cove. The uplands and highest peaks, however, were never inhabited on a permanent basis. Instead they served as hunting grounds, summer pasture for livestock, and later as a source of trees for a timber industry (Smoky Mountain Historical Society 1986:2-3; Dunn 1988).

Sevier County's population was small in these early years, but it grew steadily. In 1800 there were 3,419 inhabitants of the county. This number had increased to 6,920 by 1850, and to over 11,000 by 1870 (Ninth Census of the Population of the United States 1870:61).

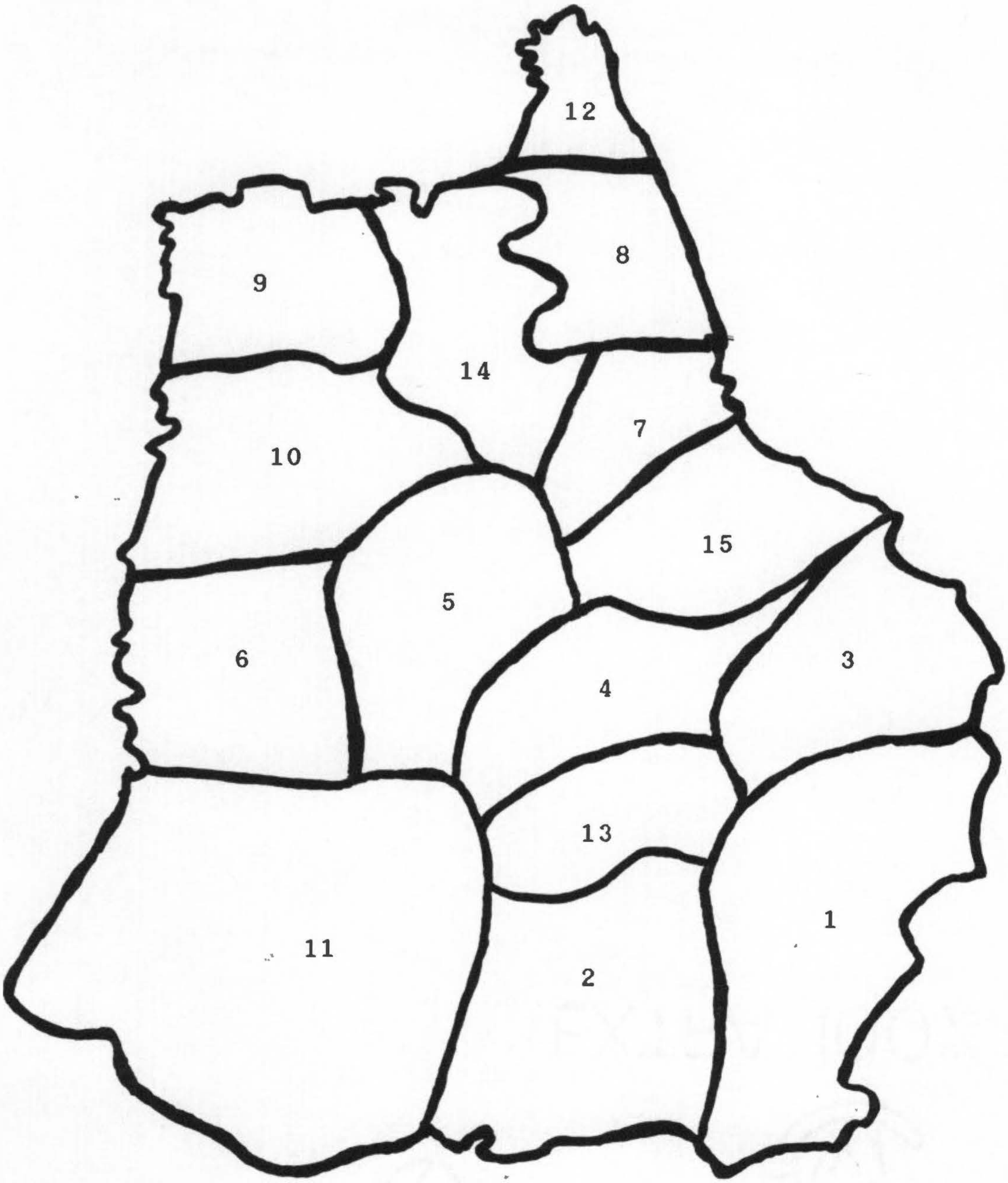
Ethnically the early pioneers were an amalgamation of British, Irish, German, and French stock. In addition, a small number of slaves and free blacks lived in the county, and some admixture probably occurred between the settlers and the aboriginal Cherokees.

These settlers were not, however, a group of unrelated individuals. Migration into this region was often a family affair--the movement of several members of an extended family. A new homesteader could usually find numerous siblings, cousins, or other kinfolk among his neighbors (Dunn 1988:8-21).

The actual extent of 19th century immigration into Sevier County is unclear. Only a small percentage of the inhabitants in 1800 were born in the region, but by 1870 over 90 percent of the population was native to the State of Tennessee. Most of the remainder came from North Carolina and Virginia, while a handful of individuals "hailed from" Alabama, Kentucky, or Georgia. Only five Sevier Countians in 1870 were foreign-born; three were English or Welsh, and two were from Ireland (Ninth Census of the Population of the United States 1870: 371).

Some writers have claimed that the area became increasingly isolated economically, politically, and socially, after the Civil War (Cummings 1988; Dunn 1988; Pearsall 1959). According to these researchers, out-migration was prevalent as people left for economically desirable jobs in the cities or for the new, unsettled lands of the American West. Few new settlers came to replace them, but the population of Sevier County still grew throughout the latter 19th century, until it reached 22,021 inhabitants at the turn of the century (Smoky Mountain Historical Society 1986:3).

This population growth led to changes in the civil districts into which Sevier County was divided (Figure 2). In 1860 there were twelve districts in the county. Each was centered around a community and division was based on population size. The majority of the inhabitants lived in the northwest and central sections of the county, while less than



1. Jones Cove	6. Wears Valley	11. Gatlinburg
2. Emerts Cove	7. Catlettsburg	12. Sinking Springs
3. Fair Garden	8. Henrys Xroads	13. Richardsons Cove
4. Harrisburg	9. Trundles Xroads	14. Boyds Creek
5. Sevierville	10. Cusicks Xroads	15. Allensville

Figure 2. Civil Districts of Sevier County: 1880

10 percent dwelled in the two most isolated, mountainous districts (Cummings 1988:59). The boundaries of all districts were very irregular, due more to the broken topography and differential distribution of citizens than to the gerrymandering common in other parts of the country.

Henrys Crossroads (District 8) and Sinking Springs (District 12) lay in the triangle north of the French Broad, while Catlettsburg (District 7) and Trundles Crossroads (District 9) lay along the other bank. These districts were completely in the northwestern section of the county--the floodplain of the French Broad.

Other districts had their northern edges in this river valley, but became more rugged toward the south. These included Sevierville (District 5), Fair Garden (District 3), and Harrisburg (District 4). Cusicks Crossroads (District 10) and Wears Valley (District 6) were on the western side of the county, while Jones Cove (District 1) lay on the eastern end and stretched south into the higher mountains.

Along the boundary with North Carolina were the two most mountainous districts, Gatlinburg (District 11) and Emerts Cove (District 2). These districts were separated from the rest of Sevier County by narrow gorges or high ridges, while towering mountains lay between them and communities across the state line.

By 1870 population growth had led to the formation of two new districts. Richardsons Cove (District 13) was fashioned from sections of Harrisburg and Emerts Cove, while Boyds Creek (District 14) was made from a section of Trundles Crossroads[1].

Four more districts were created before the turn of the century. In 1876 Allensville (District 15) was established from parts of Fair Garden, Harrisburg, and Catlettsburg, while Wears Valley was divided in the 1890's to create the district of Waldens Creek (District 16). Finally, parts of Emerts Cove and Jones Cove became the Greenbriar District (District 17) between 1890 and 1900 (Cummings 1988:59).

If the 200 percent growth of the population between 1850 and 1900 was due to natural increase rather than immigration, then an increase in homozygosity would be expected as well. Kinship in a genetically isolated population would have increased from 1850 to 1900.

Unfortunately, no systematic examination of late 18th century immigration has been undertaken. In addition, no study has been made of the degree of population subdivision within the county. There is not even an estimate of the average kinship between early Sevier Countians. If we are to recognize and understand any increase in homozygosity, then we must have objective measures of these features of the genetic and demographic structure.

---

[1] Nineteenth century court records note the creation of new districts, but do not indicate their boundaries. The details given here are based on an analysis by Cummings (1988:59), and information provided by Beulah Linn, the Sevier County Historian.

This in turn presents a problem. Investigating the genetics of a historical population is very difficult. Serogenetics and anthropometrics, widely used with living populations, cannot be applied with those long dead. Osteometric analysis is useful where skeletal samples are available, but for Appalachia no such samples exist. The major means available for exploring the genetic structure of a historic group are path analysis of genealogies, a complex and difficult method, and the analysis of surname distributions in historical records.

CHAPTER 3

Surnames and Genetics

That surnames, at least Western surnames, operate as crude genetic indicators should not be surprising. In western cultures a last name symbolizes an individual's membership in a family or lineage and, in addition to being a social designation, indicates a relationship with other members of the lineage. Because surnames are passed from parent to child, their inheritance is roughly analogous to the inheritance of sex-linked genes, and the possession of like surnames, a situation known as "surname isonomy", often indicates genetic kinship.

What is surprising is that geneticists took so long to recognize the value of using surnames to estimate the genetic parameters and structures of human populations. George Darwin in 1875 suggested that the level of inbreeding could be estimated by counting marriages between individuals with the same patrilineal surname. In 1954 Kamizaki calculated the probability of sharing a like surname for different degrees of relationship (Kamizaki 1954; Crow 1983). Twelve years later, other researchers independently worked out the same calculations as Kamizaki and were able to devise objective measures of inbreeding based on surname isonomy.

These researchers, Crow and Mange, (1965) pointed out that when mating is random, the likelihood of inheriting an identical surname from a common ancestor is proportional to the likelihood of inheriting an identical autosomal gene from that same ancestor at a 2 to 1 ratio. Brothers and sisters always have the same surname, but only inherit identical genes from the same parent 50 percent of the time. First cousins share the same name 25 percent of the time but only inherit 12.5 percent of the same genes. This ratio remains consistent as relationships become more distant and even applies in relationships such as uncle-niece and aunt-nephew.

The correlation between surname and genetic inheritance enabled Crow and Mange (1965) to develop a formula based on the proportion of isonomy between marital pairs that determines the inbreeding of offspring. If the probability of inheriting an autosomal gene from an ancestor is 1/2 the probability of inheriting a surname from that ancestor, then the proportion of homozygosity for that gene resulting from current inbreeding should be 1/4 of the proportion of isonomous marriages in the parent generation. Crow and Manges' method also distinguishes between random (F_r) and nonrandom (F_n) components. The formula is

$$F = F_n + (1 - F_n)F_r \tag{1}$$

where

$$F_n = (P - (\Sigma\, p_i q_i))/4(1 - (\Sigma\, p_i q_i)) \tag{2}$$

and

$$F_r = \Sigma(p_i q_i)/4 \tag{3}$$

In the above formulae, P equals the proportion of isonomous marriages in the marital population, while p_i equals the proportion of males with the i^{th} surname, and q_i equals the proportion of females with the i^{th} surname.

This method will be accurate only if certain assumptions about surnames and their distribution are valid. Most important is that surnames are monophyletic in origin--that people with an identical surname inherited that surname from a common (though perhaps distant) ancestor. Similarly, all individuals descended from the same patrilineal ancestor must share the same surname--changes in spelling or pronunciation must not be so great that variants of the surname are not recognizable as such. Finally, surnames must be selectively neutral--there must be no selective advantage or disadvantage to the possession of a surname or class of surnames.

Rarely if ever are these assumptions fully met in real populations. Polyphyletic origin of surnames has presented problems in many studies (Azevedo et al. 1969; Lasker 1985; Devor 1980; Jorde and Morgan 1987), and other researchers have documented the mutation of original surnames into dissimilar variants (Weiss et al. 1983). Even the selective neutrality of surnames has been challenged in some cultures (McCullough et al. 1985), and is questionable in Appalachia (Batteau 1982).

In addition, Crow and Mange's methods present certain theoretical problems. Age differences between mates and marital propinquity can skew the results, leading to overestimation of the nonrandom component of inbreeding and underestimation of the random component (Dyke et al. 1983). There is some ambiguity about the calculation of random and nonrandom components and their interpretation (Libet 1983; Crow 1983). Also, migration, especially differential migration of males and females, can systematically bias results.

If the influence of these factors is appraised and the results are considered as relative indicators rather than absolute measures, then Crow and Mange's method may still provide useful information about the genetic structure of populations. Surnames may also be used in other ways to investigate this structure.

For instance, the correlation between surname and genetic inheritance permits the estimation of kinship within or between local populations based on the commonality of surnames within or between those populations. This does not require an examination of marital isonomy, but instead surveys overall isonomy in lists of surnames such as found in census and voting records.

A method of estimating kinship from such records has been suggested by Relethford (1988). If genetic kinship is defined as "a measure of the loss of heterozygosity relative to some reference population" (Relethford 1988:476), then a coefficient of kinship (k) can be determined by the formula

$$k = I_{ij}/4 \tag{4}$$

where I_{ij} is the random isonomy between population i and population j (Relethford 1988:477). I_{ij} may be derived from the formula

$$I_{ij} = \Sigma(n_{ik}n_{jk})/N_iN_j \tag{5}$$

where n_{ik} and n_{jk} are the number of individuals with surname K in populations i and j respectively, and N_i and N_j are the sample sizes of populations i and j (Relethford 1988:478).

Within-group kinship may be estimated by formula 4 if I_{ij} is replaced with I_{ii} taken as

$$I_{ii} = \Sigma\ (n_{ik}\ (n_{ik} - 1))/N_i(N_i - 1). \tag{6}$$

Formula 4 refers to *a priori* kinship or kinship relative to an ancestral or founding population. As Relethford (1988) has pointed out, the concept of a "founding population" in a genetic sense is not valid for most human groups. Instead he suggests the use of *conditional* kinship--kinship relative to regional mean kinship--as a more meaningful indicator of genetic relationships.

Conditional kinship (k_c), according to Relethford, may be derived from surname data using the following formula

$$k_c = (I_{ij} - R)/4 - (1 - R) \qquad (7)$$

where R is the regional mean kinship taken as the weighted average of all within and among population kinships, or by applying formula 6 to the regional population as a single group (population i).

Measures of microdifferentiation can also be derived from these kinship coefficients (Relethford 1988:480-482). A measure analogous to Wright's F_{st}, (which refers to a theoretical founding population) can be calculated "as the weighted mean value of all within-group *a priori* kinship" (Relethford 1988:480) or by the formula

$$F_{st} = \Sigma\ w_i k_{ii} \qquad (8)$$

where the weighting factor (w_i) is the proportion of subgroup i in the total population or N_i/N_T, k_{ii} is the within-population kinship of group i as determined by formulae 4 and 6 above, and the summation is for all sub-groups contained in the regional or total population.

If, as Relethford suggests, measures of *conditional* kinship are usually more relevant than measures of *a priori* kinship, then the most appropriate coefficient of microdifferentiation should be calculated as

$$R_{st} = \Sigma\, w_i kc \tag{9}$$

where the weighting factor (w_i) is determined as above, and where the kinship coefficient (kc) is calculated using formula 7.

Isonomy can also be used to calculate genetic distances, both *conditional* and *a priori*, between populations. However, Relethford (1988:484) has suggested that the uncertain validity of the assumptions of surname analysis and resultant variation in its accuracy, preclude the calculation of absolute genetic distances. Instead he proposes a relative measure of genetic distance calculated by the simple formula

$$d^2 = I_{ii} + I_{jj} - 2I_{ij}. \tag{10}$$

This formula provides a comparable distance measure (d^2) between local research samples.

The relationships between several populations or subpopulations can be further explored by using formula 10 to formulate a matrix of d^2 distances. The results, though Euclidean within multidimensional space, are difficult to interpret within two- or three-dimensional theoretical models.

One way to clarify these relationships is to employ principal coordinates analysis as described by Gower (1972).

Principal coordinates analysis scales the distances in a matrix onto orthogonal variates. Although the potential number of variates or *eigenvectors* (and therefore the potential number of dimensions) that can be derived by this procedure is equal to the number of groups less one, a large portion of the information is expressed in the most important variates. Thus, two- or three-dimensional displays based on primary eigenvectors will often accurately reflect relationships with little loss of information resulting from the exclusion of less significant variates.

Another important aspect of genetic structure which can be addressed through surname analysis is the subdivision of a population into smaller breeding units or rather the degree of a population's departure from panmixia. This subdivision may be indicated by the ratio of total inbreeding to expected or random inbreeding, as determined by Crow and Mange's method. The ratio may be expressed as an index.

$$F/F_r \qquad (11)$$

If a population practices true panmixia then this index should equal approximately 1. In reality this is rarely the case. Avoidance of mating between close relatives lowers this index while other factors, such as localized mate selection, tend to increase it. The index is an indicator of the degree of nonrandom behavior, but does not distinguish the causes of such behavior.

One problem with using marital isonomy in this manner is that the sampling variance is high. Lasker (1985:24-25) has shown that because of the large number of surnames within most populations, very large samples are needed to accurately estimate marital isonomy. This is less of a problem when dealing with just the random component because the sample size in this case is $N \times (N - 1)$ where N equals the number of marriages. Therefore, it is the nonrandom component of marital isonomy that is suspect, and any measures derived from it, like F and F/F_r, must be understood in context of the sample size.

Another measure of a population's departure from panmixia (RP) has been proposed by Lasker and Kaplan (1985). This method is based not on marital isonomy but on the repetition of certain pairs of surnames in marriage records. The formula is

$$RP = \Sigma \, [S_{ij}(S_{ij} - 1)]/N\,(N-1) \tag{12}$$

in which S_{ij} is the number of couples where the male has the i^{th} surname and the female has the j^{th} surname, and where N is the number of marriages. An analytical method of determining the random expectation of RP has been worked out by Chakraborty (1986). This formula is

$$RP_r = [(1/(N-1))\ \Sigma(S_{i.}^2-(1/(N-1)))]\ x$$
$$[(1/(N-1))\ \Sigma(S_{.j}^2-\ (1/(N-1)))] \qquad (13)$$

where $S_{i.}$ is the number of males with the i^{th} surname and $S_{.j}$ is the number of females with the j^{th} surname.

An index similar to formula 8 can be calculated from RP and RP_r.

$$RP/RP_r \qquad (14)$$

This index, like F/F_r is a measure of the ratio of total repeated marital pairs to expected repetitions, and as such, is a measure of the subdivision of a population. While both indices are related, they are not equivalent. Certain mating patterns will affect the two differently. Avoidance of isonomous matings and extended reciprocal mate exchanges between lineages will artificially deflate F/F_r without systematically affecting RP/RP_r. Because of this, and the much smaller sampling variance of RP, the latter index may be more dependable in revealing genetic subdivision.

Lasker (1988) has also used RP with individual lineages (referred to here as RP_l), in which all individuals possessing a certain surname are considered a lineage, and the coefficient is based on their spouses' surnames. In this method sexes are pooled, and isonomous marriages are not included in the analysis. The formula used is identical to *RP* as given previously, but S_{ij} is redefined as the number of couples where one individual has the i^{th} or lineage surname and his or her spouse has the j^{th} surname, and where N refers to the number of marriages in the lineage sample.

Comparing the genetic structures of very different populations by comparing coefficients and measurements derived from surnames is a technique fraught with the potential for error. It can be like comparing proverbial apples and oranges because different nations and ethnic groups have used surnames in different ways.

In continental Europe, for instance, the use of patrilineal surnames began around 1000, probably in Normandy. Their use, as well as the surnames themselves, spread to other parts of the continent and, with the Norman Conquest, to the British Isles. In Scotland and other parts of Celtic Britain clan names already existed, but clans and lineages were quite different. As one early researcher explained:

> Scottish surnames are doubtless a difficult subject to deal with, and this principally by reason of the system of clanship so long prevalent in that kingdom. In Scotland whoever joined a particular clan, no matter what his position or descent, assumed the surname of his chief, and this was accepted as an act of loyalty. In England, had any retainer of a feudal baron presumed to do such a thing he would have found himself at the bottom of the deepest dungeon of the castle! (Lower 1988:xxi)

Clearly "MacGreogor" has different genetic implications than does "Windsor".

Other European countries have used different schemes. In Scandinavia surnames were once patronymic--derived from the father's first name. If a boy named Knute had a father Soren, a grandfather Peder and a great-grandfather Nils, then the names of the boy, father, and grandfather would have been Knute Sorenson, Soren Pederson, and Peder Nilson, respectively. Although patrilineal inheritance of surnames has now been adopted in Scandinavia, the earlier system resulted in a very polyphyletic modern assortment of surnames.

In non-western cultures histories of surname usage are very diverse as well. The surname system in China is the most ancient in the world dating to the Han Dynasty around the time of Christ (Lasker 1985:39), while the aborigines of Taiwan only adopted surnames after World War II (Chen and Cavalli-Sforza 1983:367). Somewhere in the middle of these extremes is Japan, where of the approximately 120,000 extant Japanese surnames only about 30,000 predate an 1875 government declaration making surnames mandatory (Yasuda 1983:265).

The comparison of isonomy coefficients between cultures with diverse surname systems and histories has little value even in theory. Comparisons between groups of similar ethnicity is more appropriate, especially when the similarities are strong. Sevier County, for example, with its population of British and Northern European descendants is comparable only with populations of similar descent.

The indexes F/F_r and RP/RP_r avoid the problem of different surname systems because they examine the relationship between random and nonrandom mating behavior, rather than the distribution of alleles. These indexes can be used rather freely to compare population subdivision in communities of disparate ethnicity.

Obviously surname analysis, if done properly and kept in perspective, can be a valuable and versatile tool for doing historical genetic research. Because of the nature of the records available, it is also the only feasible means of exploring the genetics of late 19th century Sevier County. Therefore this project has employed surname analysis as its method of investigation.

CHAPTER 4

Materials and Methods

The settlement of Sevier County probably began "earlier than some history books record (Smoky Mountain Historical Society 1984:V)". Because few records of the county's early years survive, the names of the first settlers and the date of their arrival are unknown. Censuses, land grants and church records from the 19th century do exist, but historical documents dating to the late 18th century are almost unknown. Even records from the first half of the 19th century are scarce--a fire in March of 1856 destroyed the Sevier County Courthouse and with it all deeds, wills, and court records.

Federal census records for this area began in 1790, when it was still a part of North Carolina. The population count and its breakdown by sex, age, and race are known from later references, but the names of these then North Carolinians are unavailable. Later censuses do record the surnames of heads of households and also the sex and approximate age of each individual within the households. Beginning with the census of 1850, the names, race, sex, and ages of all the inhabitants of the county were recorded. This census and those of 1860, 1870, and 1880 have been transcribed by local historians and genealogists. The transcriptions include an index of the names of all heads of households and indicate the particular districts where each family resided.

Because marriage licenses were kept in bound volumes in the county courthouse, they too were destroyed in the 1856 fire. Extant marriage records begin in April of that year with a license issued to P.M. Atchley and Margaret E. Thomas. The records are largely complete for the remainder of the 19th century except for the civil war years, 1861 to 1865, when the difficulty of traveling to the county seat apparently led many couples to wed without the sanction of county authorities. After the fall of the Confederacy, many of these war-time marriages were recorded at the courthouse and were backdated with the actual marriage date.

Two separate entries were required to correctly record a marriage. First the date of license issuance was recorded on a page of the marriage book. This entry included the names of the betrothed couple and the signatures of two witnesses. Next, the date of the wedding or "solemnizing" was entered on the bottom of the same page by the official who performed the ceremony, a Justice of the Peace or occasionally a minister. Often the wedding was performed the same day, and in almost every case took place within a few days.

When a wedding was called off, the license was returned without this second entry. Certain county clerks were less than diligent in this process, and the second entry was sometimes neglected for consummated marriages. A strict reading of the marriage books, as these bound volumes are called, would indicate that in particular years no couples made the stretch from courthouse to alter. During the terms of more assiduous clerks, returned or incomplete licenses were infrequent, even rare.

The number of marriages increased along with the population throughout the late 19th century. From 1856 to 1865 only 818 marriages were registered in the county, but that number grew to 1,192 during the next ten years and to 1,344 between the years 1876 and 1885. The increase was even greater from 1886 to 1895 when 1,910 licenses were issued. Then, in the decade straddling the turn of the century, 1,827 "licensed" couples walked down the aisle in Sevier County.

If each different spelling in the marriage records was in fact a separate surname, then the people of Sevier County would have been among the most genealogically diverse in history. Some of this apparent diversity, however, was a result of the high illiteracy rate in the county. In 1850 there were 1,005 White adults in Sevier County who could not read or write--40% of the White population over 20 years of age (United States Census of 1850:583, table IX). This illiteracy is also evident in the marriage records where "His Mark 'X'" often appears in longhand beside an individual's name. The full name in such cases was written out by the county clerk, and this practice may have contributed to the mutation of some surnames.

Often the connection between variants was obvious, as with Partin, Parton, and Purtin. In other cases the differences in spelling left the relationships cloudy--Ownby became Owenby and Ownsby, but did it also become Honby? Was McCulla the same as McCullough? And who were kinfolk among the Laymans, the Lehmans, the Lemans and the Lemons?

When the apparently related variants are combined, the number of different surnames in the marriage books from 1856 to 1906 exceeds 1,000. Included among these are patronymic surnames such as Williams, Johnson, and Robertson, and occupational surnames like Smith, Clark (clerk) and Chandler (candlemaker). Descriptive surnames like Reed (red), Sharp (from a disposition) and Moore (from residence near a boggy heath) are as common as names derived from animals: Bird, Fox, Wolf. Some names are geographical, referring to the area of a surnames origin; Breeden is derived from Bredon, a parish in Worchester, while "Underwood" is a township in Derbyshire (Lower 1988).

The majority of these surnames originated in the British Isles--of the 100 most common names in the marriage records 90% were listed in A Dictionary of the Family Names of the United Kingdom published in 1860 (Lower 1988). Some, such as Hurst, Webb, and Walker, are Anglo-Saxon names. Some, like Allen, Henry, and Lowe, date to the Norman conquest. Still others have ancient Celtic roots--Reagan, McMahan, Lewelling.

Names originating in continental Europe are present, though not in great numbers. Of the 100 most common names only the German "Shultz" is clearly not British. French names like Lafollette and Serarat, though relatively rare, do occur. The ethnic homogeneity evidenced by the surnames, however, may be in part an illusion.

Durwood Dunn has shown how the Pennsylvania German settlers of Cades Cove, a mountain valley in neighboring Blount County, anglicized their surnames; Koebel became Cable for instance (Dunn 1983). This may have occurred in Sevier County as well, where many of the early settlers came down the Great Valley from Pennsylvania. The Fox family, ostensibly English, descended from Pennsylvanian ancestors, who left at least some family documents written in German. The Clabough, Parrott, and Romines families, on the other hand, may have descended from French ancestors.

Clerks who recorded the surnames likely played a role in their Anglicization. Unacquainted with umlauts and cedillas, they may have written foreign names as more familiar English forms. Also, because clerks tended to spell phonetically, it is difficult to differentiate in the historical documents between British surnames and ones from the European continent.

While these problems present questions about the ethnic homogeneity of the Sevier Countians, they also introduce uncertainty concerning the monophyletic origins of surnames. Still, if this uncertainty is acknowledged and is considered when interpreting the data, then surname analysis is a credible method of examining the genetic structure of 19th century Sevier County.

Marital isonomy was examined for 5 separate decades: 1856-1865, 1866-1875, 1876-1885, 1886-1895, and 1896-1905. Crow and Mange's method was used, first calculating random and nonrandom components (formulas 3 and 2, respectively) and then solving for the inbreeding coefficient, F. (formula 1). The index F/F_r was then used to calculate the ratio of total inbreeding to its random component.

A priori kinship coefficients (k) were also calculated from random isonomy within the marriage records. Male and female within-sex kinship was computed, as was between-sex kinship for each decade.

Lasker and Kaplan's repeated pairs analysis, RP, was applied to these 5 samples (formula 12), and the random component, RP_r, was calculated following Chakraborty's formula (formula 13). The departure of RP from its random expectation was determined using the index RP/RP_r. In addition, lineage-specific RP_1 was calculated for the twenty-five most common surnames in these marriage records.

Within- and between-district isonomy was calculated from indexes of the 1860, 1870, and 1880 census transcriptions using formulas 5 and 6. From the resultant coefficients, *a priori* kinship (k) and *conditional* kinship (kc), were determined between and within all civil districts for each census period. Within-district kinship was used to calculate the *a priori* and *conditional* measures of microdifferentiation, F_{st} and R_{st}.

Finally, relative genetic distances (d^2) between all districts were computed, and matrices of such distances were formed for each census year. Principal coordinates analysis as described by Gower (1972) was applied to these matrices and the results were plotted on two-dimensional graphs.

All calculations were performed with a microcomputer using a variety of software packages and programming languages. The principal coordinates analysis of the d^2 matrices was achieved using a BASIC program written by Dr. Richard Jantz of the University of Tennessee, Knoxville.

CHAPTER 5

Surnames and Structure: The Marriage Records:

The results of marital isonomy analysis for the fifty year period, 1856-1905, are given in Table I. The inbreeding coefficient, F, decreased in the years following the Civil War, increased again between 1876 and 1895, and then dropped slightly in the next decade. The nonrandom component, F_n, followed the same pattern, as reflected in both its value and the value of the index F/F_r. The random component, however, decreased from the Civil War years through 1876-1885.

TABLE I

Results of the Marital Isonomy Analysis
Sevier County, Tennessee: 1856-1905

Years	F	F_n	F_r	F/F_r	No. of Marriages
1856-1906	.0065	.0056	.0012	5.42	7,091
1856-1865	.0064	.0052	.0012	5.33	818
1866-1875	.0055	.0043	.0011	5.00	1,192
1876-1885	.0063	.0053	.0010	6.30	1,344
1886-1895	.0076	.0064	.0012	6.33	1,910
1896-1905	.0071	.0058	.0014	5.07	1,827

Results of repeated-pairs analysis are presented in Table II. All measures decreased in the postbellum years; RP and RP_r declined through 1876-1885, while RP/RP_r continued to decrease into the 20th century. Because the distribution of these coefficients is asymmetrical and non-normal, it is difficult to assess the statistical significance of these temporal changes.

TABLE II

Results of the Repeated Pairs Analysis
Sevier County, Tennessee: 1856-1905

Years	*RP*	RP_r	RP/RP_r	No. of Marriages
1856-1865	.000180	.000025	7.20	818
1866-1875	.000135	.000022	6.14	1,192
1876-1885	.000103	.000017	6.06	1,344
1886-1895	.000121	.000024	5.04	1,910
1896-1905	.000112	.000030	3.73	1,827

Finally, the results of between-sex and within-sex kinship analysis are presented in Table III. Between-sex *a priori* kinship (equivalent to F_r) decreased until 1886-1895, then increased through the turn of the century. Within-sex kinship followed the same pattern for both genders, but males showed slightly less within-sex kinship than females during the last four decades studied.

The measures of random isonomy, F_r, RP_r, and k, indicate that average kinship in Sevier County decreased immediately after the Civil War. Net inbreeding was apparently not occurring; on the contrary, the coefficients imply that the population as a whole was experiencing a period of decreasing genetic homogeneity.

TABLE III

Results of *A Priori* Kinship Analysis
Sevier County Marriage Records: 1856-1905

Years	Male-to-Female *k*	Male-to-Male *k*	Female-to-Female *k*
1856-1865	.00123	.00157	.00153
1866-1875	.00114	.00131	.00144
1876-1885	.00105	.00120	.00125
1886-1895	.00124	.00133	.00141
1896-1905	.00139	.00142	.00161

Between 1876-1885 and 1896-1905, however, random isonomy increased, indicating that the breeding population became more isolated genetically and that inbreeding intensified. By 1896-1905, all random-isonomy based coefficients (except male within-sex kinship) had climbed to levels exceeding those of 1856-1865.

The divergence of the inbreeding coefficient F from this pattern is a result of the influence of its non-random component, F_n, which has a large variance. *RP*, which has a non-random component with a smaller variance, conforms more closely with the pattern of the random components. Therefore the divergence of F should not be given excessive weight.

The decline of average kinship from 1866 to 1885 is contrary to the idea of a postbellum increase in the genetic isolation of Sevier County, but this does not necessarily mean that "new immigrants" entered the county. The decrease in random isonomy through 1885 may have resulted from a "coming of age" of the children of an earlier wave of immigrants. Conversely, the increase in isonomy after 1886 does not establish that the county was then more isolated. Instead this may be an effect of the attainment of majority by children of the more closely-related 1856-1865 generation. The fluctuation in random isonomy could be a generational phenomenon rather than reflecting contemporaneous migration or isolation.

Whatever the cause of this fluctuation, the genetic implications are not changed. If the differences are truly significant and isonomy accurately reflects genetic structure, then Sevier County's *breeding population* was less genetically homogenous during 1866-1885 than during 1886-1905.

An interesting aspect of the kinship coefficients is the greater within-sex kinship for females than for males in the last forty years of this study. This could be due to a greater mobility for males in the region. If 19th century Appalachia males relocated more often than females, then local male-to-male kinship would be less than female-to-female kinship. Still, differences in these records are slight and it is unwise to generalize from them without further study.

These kinship estimates range from .0010 to .0016, which is a level of kinship between that of third and fourth cousins--not a particularly close relationship. Compared with other Appalachian populations, inbreeding and kinship in Sevier County was not high. As Table IV illustrates, the inbreeding coefficient in Old Morgan County, Kentucky was .0115 in 1850, and .0134 in 1870. Unlike Sevier County, F here decreased through the remainder of the 1800's and into the next century. F_r in Old Morgan County was also greater than its Sevier County counterpart and remained relatively stable through the 19th century. Both F_r and F_n, however, decreased starting around the turn of the 20th century (Tincher 1980).

The only other studies of Appalachia that employ surname isonomy are of Hancock County where three 20th century groups have been sampled. The Melungeons, a triracial isolate with somewhat mysterious origins, were studied by Politzer and Brown (1969) in the 1960's. Inbreeding in this group was augmented, not only by geographical isolation, but by social isolation from the other inhabitants of the region who viewed the Melungeons as a separate people of somewhat lower status.

The coefficient of inbreeding for the Melungeons was two to three times that found in 19th century Sevier County and almost five times that found in Old Morgan County in 1970 (Table IV). The nonrandom component of inbreeding was

TABLE IV

Comparison of Inbreeding Coefficients of Appalachian Populations

Population	F	F_n	F_r	F/F_r	N
Sevier Co. Tennessee					
1856-1865	.0064	.0052	.0012	5.33	818
1866-1875	.0055	.0043	.0011	5.00	1,192
1876-1885	.0063	.0053	.0010	6.30	1,344
1886-1895	.0076	.0064	.0012	6.33	1,910
1896-1905	.0071	.0058	.0014	5.07	1,827
Old Morgan Co. Kentucky[A]					
1850	.0115	.0085	.0030	3.83	88
1870	.0134	.0105	.0029	4.62	113
1890	.0098	.0067	.0031	3.16	206
1910	.0074	.0055	.0019	3.89	409
1930	.0054	.0040	.0015	3.60	414
1950	.0049	.0035	.0015	3.26	509
1970	.0023	.0010	.0013	1.77	431
Hancock Co. Tennessee					
Grandparents[B]	.0456	.0395	.0064	7.13	100
Parents[B]	.0099	.0041	.0058	1.71	256
Melungeons[C]	.0102	-.0051	.0152	0.67	478

[A]Tincher 1980. The coefficients are based on surname isonomy in the marriage records of the extant counties, of Morgan, Wolfe, Elliott, and Magoffin.

[B]Kirkland and Jantz 1977. These coefficients are based on marital isonomy of the parents and grandparents of Hancock County school children in 1972 and 1974.

[C]Politzer and Brown 1969. These coefficients are based on marital isonomy, but include comparisons of various relationships, such as "wife and husband's father", "husband and wife's mother, etc. The Melungeons are a triracial isolate population.

negative, indicating an approximation of panmixia within the Melungeon community except for the avoidance of close-kin mating. The random component (or average kinship), however, was quite high at .0152, approximately the kinship of second cousins.

A general study of the Hancock County population was done by Kirkland and Jantz in the 1970's. Information on the families of school children was used to estimate inbreeding and kinship for two generations. *F* for the grandparent's generation was extremely high at .0456, but for the parent's generation *F* was .0099, close to the estimate of .0102 for contemporaneous Melungeons. Kinship (random isonomy) estimates for both the grandparent's generation and the parent's generation are similar--.0064 and .0058, respectively.

These studies all suffer the limitation of small sample sizes. Samples from Old Morgan County range from 88 couples in 1850 to 509 couples in 1950. The Kirkland and Jantz study employed samples of 100 pairs of grandparents and 256 pairs of parents. The Melungeon sample used 472 comparisons between various relations, but only 72 marital couples were surveyed. Thus, the estimates of *F* are not highly reliable, though the estimates of the non-random component (*a priori* kinship), are more credible. Table V compares estimates of kinship for various North American populations of West European descent.

TABLE V

Comparison of *A Priori* Kinship Coefficients of Various North American Populations of West European Descent

Population	Kinship Coefficient k	Source
Sevier Co. Tennessee		
1856 - 1865	.00123	Marriage Records
1866 - 1875	.00114	" " " " "
1876 - 1885	.00105	" " " " "
1886 - 1895	.00124	" " " " "
1896 - 1905	.00139	" " " " "
1860	.00111	1860 Federal Census
1870	.00107	1870 " " " " "
1880	.00107	1880 " " " " "
Hutterites	.04450	Crow and Mange 1965
Utah Mormons		Jorde and Morgan 1987
1820 - 1829	.00025	Table 2
1860 - 1869	.00047	" " "
1900 - 1909	.00041	" " "
Deerfield, Massachusetts		Swedlund and Boyce 1983
1770 - 1789	.00262	Table 2
1790 - 1809	.00183	" " "
1810 - 1829	.00208	" " "
1830 - 1849	.00203	" " "
Ramea Island, Newfoundland	.00640	Devor et. al. 1984
Fogo Island, Newfoundland	.00420	Lasker 1985:45
New York Colony Before 1784	.00019	Lasker 1985:43
Vinalhaven Island, Maine	.00205	Sorg 1983
Old Colony Mennonites		Allen 1988
Before 1890	.00860	Table 2
1890 - 1921	.00870	" " "
1922 - 1930	.00850	" " "
1931 - 1950	.00960	" " "

Also included in the Table are estimates of Sevier County *a priori* kinship based on federal census records.

Table V reveals that Sevier County's kinship estimates are intermediate compared with estimates from other populations. Average kinship was certainly not in the range seen in true breeding isolates like the Hutterite or Mennonite populations, and two island populations from Newfoundland have produced coefficients much larger than those found in Sevier County. Other studies such as those of Vinalhaven, Maine, and Deerfield, Massachusetts found more comparable levels of kinship.

On the other hand, some populations have shown less mean regional kinship than Sevier County. Coefficients for the Utah Mormons, a genealogically diverse population, ranged from .00025 to .00047, and these may be overestimates because Mormon surnames tend to be polyphyletic in origin. An analysis of 18th century marriage records in New York found an average kinship of only .00019, just a fraction of the levels found in 19th century Sevier County.

There is strong congruence between kinship derived from Sevier County marriage records and kinship derived from the federal censuses. The marriage record estimates are higher, but only slightly. If all estimates are combined then mean kinship is .00116 with a standard deviation of .00011. Thus, to the extent isonomy reflects genetics, these estimates are accurate to the third or fourth decimal place.

The evidence indicates that regional mean kinship for Sevier County was not particularly high, but this does not rule out the existence of more interrelated subdivisions within the population. Certainly the ratio of random and non-random inbreeding suggests this, as does the relationship between RP and RP_r. Perhaps isolated, endogamous clusters existed in Sevier County.

RP, unlike F, does not have a straight-forward genetic interpretation. It is influenced by the number and frequency of surnames, and therefore is correlated with isonomy. But RP's primary significance as a measure of population subdivision lies in terms of the ratio of its random and nonrandom components. The index RP/RP_r illustrates this, and Table VI compares RP/RP_r in several populations.

Sevier County had a RP/RP_r ratio of 7.20 in 1856-1865. By 1896-1905 this ratio had decreased to 3.73. RP_r, which was only 14 percent of the total proportion of repeated-pairs during the first decade, increased to 27 percent by 1896-1905. These changes present clear evidence of a decrease in population subdivision in the late 19th century.

Still, breeding subdivision in Sevier County was high throughout this period compared with subdivision in other populations, especially in isolated groups such as on St. Barts and Sanday Island (Table VI). These island populations approach panmixia with RP_r comprising 98 percent and 86 percent of total RP for the samples.

TABLE VI

Comparison of Lasker and Kaplan's
Repeated Pairs Analysis for Various Populations

Population	$RP \times 10^6$	$RP_r \times 10^6$	RP/RP_r	% of Random
Sevier County, Tennessee				
1856-1865	180	25	7.20	14%
1866-1875	135	22	6.14	16%
1876-1885	103	17	6.06	16%
1886-1895	121	24	5.04	21%
1896-1905	112	30	3.73	27%
Sanday, Orkney Islands[A]	728	623	1.17	86%
Paracho, Mexico[B]	176	132	1.33	75%
Reading, England[C]	3.1	1	3.10	32%
St. Barts, Fr. West Indies[D]	7700	7550	1.02	98%
Fogo Island, Newfoundland[E]	323	117	2.76	36%

[A]Mascie-Taylor, Lasker and Boyce 1987. The random component was calculated following Chakraborty (1986).

[B]Lasker and Kaplan 1985. The random component was calculated by computer simulation.

[C]Lasker, Mascie-Taylor and Coleman 1986. The random component was calculated following Chakraborty (1988).

[D]James, Lasker and Morrell 1986. The random component was calculated following Chakraborty (1988).

[E]Koertvelessy, Crawford, Huntsman, Collins, Kelping and Uttley 1988. The random component was calculated following Chakraborty (1988).

Only two groups approach the level of breeding subdivision seen in Sevier County. Fogo Island, off the coast of Newfoundland, had a RP/RP_r index of 2.76. This high ratio results partially from the division of the population along religious lines, and if the marriage records are separated into Anglican, Catholic and Protestant samples, then the index drops to 2.02, 2.17, and 2.23, respectively.

The other study that found a large difference between total and random *RP* was of Reading, England, a modern industrial city. The dissimilarity here may have been due to the small number of repeated pairs expected at random. 2,377 of the 2,392 marital pairs studied were unique, thus the degree of population subdivision indicated by RP/RP_r may be inflated. Given the nature of industrial cities, some of the indicated subdivision may reflect class or ethnic divisions.

It is unlikely that any great ethnic or religious differences existed in 19th century Sevier County. Social status, however, has been shown to affect mate choice in other Appalachian populations (Batteau 1982) and a connection between church membership and family affiliation has been noted in at least one Tennessee mountain community (Bryant 1981). Possibly the nonrandom nature of marital behavior in Sevier County is influenced by religious or class factors. Another element in this "nonrandomness" may have been clan-like behavior among certain lineages. If so, then this should be reflected in the RP_l coefficients for those lineages.

Table VII gives RP_1 for the 25 most common surnames in the marriage records with all isonomous pairings excluded from the analysis. There is great variation from surname to surname, with RP_1 ranging from .00376 to .13829. Because RP for a one-surname lineage is equal to the coefficient of kinship (k or F_r) of spouses marrying into that lineage (Lasker 1988:2), this variation may be genetically significant. If the measurements are accurate and not solely the result of high variance or sampling error, then mates for some lineages may be more genetically homogeneous than the mates of other lineages. Mate choice for individuals with certain surnames (such as surnames 5, 10, 11 and 13 in Table VII) may have been limited by class, religion, or by the practice of clan endogamy.

Other factors may have affected mate choice. What, for instance, was the role of geography in population subdivision? If terrain somehow restricted the flow of genes and people, then this should be reflected in the distribution of surnames across the landscape. Because the marriage records were not separated by any geographic criteria, it will be necessary to turn elsewhere to examine this possibility.

TABLE VII

Repeated Pairs Analysis
for the 25 Most Common Surnames in
Sevier County Marriage Records: 1856-1905

Surname	Total No. of Individuals/Marriages	RP_1
1	351	.02270
2	226	.00790
3	219	.01152
4	189	.00589
5	177	.13829
6	170	.03627
7	167	.02836
8	159	.02589
9	158	.01048
10	147	.04102
11	146	.03544
12	137	.00376
13	133	.04276
14	128	.02362
15	127	.02181
16	120	.01275
17	119	.00707
18	117	.01083
19	111	.00623
20	104	.00803
21	101	.01869
22	99	.01075
23	99	.00795
24	93	.00536
25	91	.00635

CHAPTER 6

Surnames and Geography: The Census Records

Because Sevier County census records were divided into civil districts, these records were useful for examining the geographical distribution of surnames and by analogy, geography's influence on genetic structure. Kinship estimates derived from these records were used to assess the effect of geography on population microdifferentiation (F_{st} and R_{st}), within-district homozygosity (k and k_C), and genetic distances between districts (d^2).

F_{st} and R_{st} for 1860 are presented in Table VIII, along with estimates of *a priori* and *conditional* within-group kinship. Mean group kinship was greatest in Gatlinburg, the most isolated district, where k equaled .0151 or approximately the kinship of second cousins. Two other mountain districts, Emerts Cove and Jones Cove, had kinship estimates of .0079 and .0074 respectively. Low mean kinship (k = .0025 to .0026) existed in Harrisburg and Fair Garden in the center of the county, and in Trundles Crossroads on the south bank of the French Broad. North of this river, in Henrys Crossroads and Sinking Springs, kinship coefficients were intermediate.

All districts had positive values of conditional kinship indicating that within-district kinship was higher than mean regional kinship in every instance. This was also reflected in the related measures of microdifferentiation.

TABLE VIII

Within-District Kinship
Sevier County: 1860

District (No., Name)	*k*	*kc*	*N*
1. Jones Cove	.0074	.0063	171
2. Emerts Cove	.0079	.0068	101
3. Fair Garden	.0026	.0015	228
4. Harrisburg	.0025	.0014	216
5. Sevierville	.0037	.0026	186
6. Wears Valley	.0032	.0021	146
7. Catlettsburg	.0041	.0030	179
8. Henrys X-roads	.0059	.0048	140
9. Trundles X-roads	.0026	.0015	192
10. Cusicks X-roads	.0037	.0026	154
11. Gatlinburg	.0151	.0140	95
12. Sinking Springs	.0052	.0041	86
	F_{st} = .0047	R_{st} = .0036	Total = 1,894

Table IX summarizes these coefficients for 1870. Although mean regional kinship fell slightly, within-district kinship increased overall as reflected by increases in both F_{st} and R_{st}. Only in Sevierville, Catlettsburg, and Henrys Crossroads did within-district kinship decrease. Kinship in the mountainous regions rose only modestly, except in Emerts Cove where mean kinship increased dramatically to .0133. This may have been the result of the formation of a 13th district, Richardsons Cove, partly from the northeastern section of Emerts Cove. With newly-restricted boundaries, Emerts Cove in 1870 fell completely within the most mountainous section

TABLE IX

Within-District Kinship
Sevier County: 1870

District (No., Name)	*k*	*kc*	*N*
1. Jones Cove	.0084	.0073	165
2. Emerts Cove	.0133	.0122	90
3. Fair Garden	.0038	.0027	205
4. Harrisburg	.0033	.0023	151
5. Sevierville	.0026	.0015	165
6. Wears Valley	.0089	.0079	137
7. Catlettsburg	.0032	.0022	159
8. Henrys X-roads	.0043	.0032	172
9. Trundles X-roads	.0029	.0019	112
10. Cusicks X-roads	.0029	.0019	158
11. Gatlinburg	.0183	.0173	103
12. Sinking Springs	.0057	.0047	82
13. Richardsons Cove	.0068	.0057	88
14. Boyds Creek	.0058	.0048	102
	F_{st}= .0059	R_{st} = .0048	Total = 1,889

of the county. Mean group kinship was still highest in Gatlinburg, and conditional within-group kinship was again positive for all districts.

Kinship estimates for 1880 are given in Table X. Only Fair Garden, Jones Cove and Catlettsburg had increases in *k* and *kc*, while these measures decreased in all other districts. Isonomy was again greatest in the mountains and all conditional kinship coefficients were positive.

Estimates of kinship such as these are problematic because the statistical significance of any one difference is difficult to test. Still, certain patterns were consistently evident in the data. Most notable was the high within-group

TABLE X

Within-District Kinship
Sevier County: 1880

District (No., Name)	k	k_c	N
1. Jones Cove	.0096	.0085	284
2. Emerts Cove	.0074	.0063	207
3. Fair Garden	.0044	.0034	269
4. Harrisburg	.0026	.0015	207
5. Sevierville	.0024	.0013	323
6. Wears Valley	.0027	.0016	248
7. Catlettsburg	.0040	.0029	230
8. Henrys X-roads	.0040	.0029	243
9. Trundles X-roads	.0019	.0008	223
10. Cusicks X-roads	.0025	.0014	282
11. Gatlinburg	.0164	.0153	211
12. Sinking Springs	.0043	.0032	158
13. Richardsons Cove	.0055	.0045	148
14. Boyds Creek	.0031	.0021	190
15. Allensville	.0043	.0033	123
	F_{st} = .0049	R_{st} = .0039	Total = 3,346

kinship occurring among the mountain dwellers. Gatlinburg was uniformly the most isonomous, and Emerts and Jones Coves also had high kinship scores compared with the low-lying districts. If these surnames truly reflect genetics, then the mountainous terrain was influencing genetic structure and somehow inhibiting gene flow.

But even in the low-lying districts, gene distribution was apparently not uniform. The positive values of all conditional within-district kinship coefficients indicate some localization of gene pools, and reveal the impact of geography on surname distribution and genetic structure.

F_{st} and R_{st} are derived from these kinship estimates and provide quantitative measures of the genetic segregation of districts. F_{st}, a gauge of differentiation from a common, homogeneous, ancestral population, is not the best measure of variance among groups not descended from a distinct founder population. As calculated here, F_{st} is the weighted average *a priori* kinship for all districts. R_{st} is the average weighted *conditional* kinship, but it is also a measure of the populations' variance around a contemporary regional mean. Because there was no distinct parental population for Sevier County, at least not in the genetic sense, R_{st} is the most appropriate indicator of microdifferentiation.

Both F_{st} and R_{st} increased in Sevier County in 1870, and then decreased in 1880 to approximately 1860 levels. Because these estimates are based on lists of heads of households, this substantial shift in kinship over such a short period is not credible. Heads of households should be a relatively stable population, and a generational fluctuation, as perhaps is seen in the marriage records, would be expressed only over an interval considerably greater than a decade.

The explanation for this fluctuation may lie in the unreliability of the 1870 census. In the postbellum years there was much distrust of the federal government throughout the South, and census-takers sometimes did not enjoy the cooperation of certain segments of the public. Estimates derived from this census must therefore be viewed with skepticism, since this lack of cooperation may have systematically biased data collection. In addition, errors could have occurred during transcription of the census.

Still, mean regional kinship in 1870 was congruent with that seen in 1860 and 1880. Also the proportion of F_{st} due to R_{st} was high for each decade, suggesting a genetic subdivision of the population that mirrored somewhat the geo-political subdivision of the county.

F_{st} and R_{st}, though subject to the same limitations as all isonomy-based estimates, can be used to compare the degree of microdifferentiation between population sets. Table XI contrasts these coefficients for 19th century Sevier County with values recorded for Amish churches in Pennsylvania, Irish isolates, and early 19th century Massachusetts Townships.

Average within-group *a priori* kinship was much lower for Sevier County than for the Amish sample or the Irish isolates, but was higher than levels observed in 19th century Massachusetts townships. R_{st} followed this same pattern. Like isonomy in the marriage records, which indicated an intermediate level of regional kinship, estimates of microdifferentiation for 19th century Sevier County fall between measures of true isolates and those reported for less segregated populations.

Discussion thus far has focused on within-group kinship in its various expressions, but the relationship between

TABLE XI

Comparison of Coefficients of Microdifferentiation from Various Populations

Population	F_{st}	R_{st}	No. of Groups	Source
Sevier County				
1860	.0047	.0036	12	Census Records
1870	.0059	.0048	14	" " " " "
1880	.0049	.0039	15	" " " " "
Nebraska Amish, Pennsylvania	.1014	.0107	3	Hurd 1983[A]
Irish Isolates 1890's	.0117	.0087	7	Relethford 1982[A]
Massachusetts				
1800-1809	.0023	.0012	4	Relethford 1988
1820-1829	.0021	.0011	4	" " " " "
1840-1849	.0015	.0006	4	" " " " "

[A]These are the original studies. The calculations for F_{st} and R_{st} are given by Relethford (1988:486, Table 1)

groups is also an important element of genetic structure. Relationships between groups can be expressed by measures of genetic distance which "provide information regarding the dissimilarity of populations in unit free terms" (Relethford 1988). The coefficient d^2, described in Chapter III, depicts between-group relationships on a relative scale, with d^2 increasing as genetic distance increases--the greater the coefficient's value, the more genetically dissimilar the two populations. Census information was used to calculate d^2 matrices for Sevier County for each census year. These matrices are included in Appendix 1.

Although each estimate of genetic distance expresses a relationship between two districts, interpreting an entire distance matrix can be very difficult. One method of summarizing this information is to average distances between one group and all other groups. This has been done for each district and the results are presented in Table XII. In addition to mean genetic distances by district for 1860, 1870, and 1880, each district's ranking, from most to least isolated, is included for each census year.

Average genetic distances mirror the pattern seen in the related measures of within-group kinship. Means were greatest for the mountain districts. Gatlinburg had the largest average genetic distance in each census year, followed by Jones Cove and Emerts Cove, both with an average rank of 3. The smallest means occurred in the center of the county, in Harrisburg and Sevierville, with average ranks of 13 and 12 respectively, and in Trundles Crossroads with an average rank of 12.

Another method of displaying the information contained in a d^2 matrix involves deriving a set of coordinates on orthogonal variates for each group through principal coordinates analysis. Graphing the coordinates for the primary variates creates a visual representation of

TABLE XII

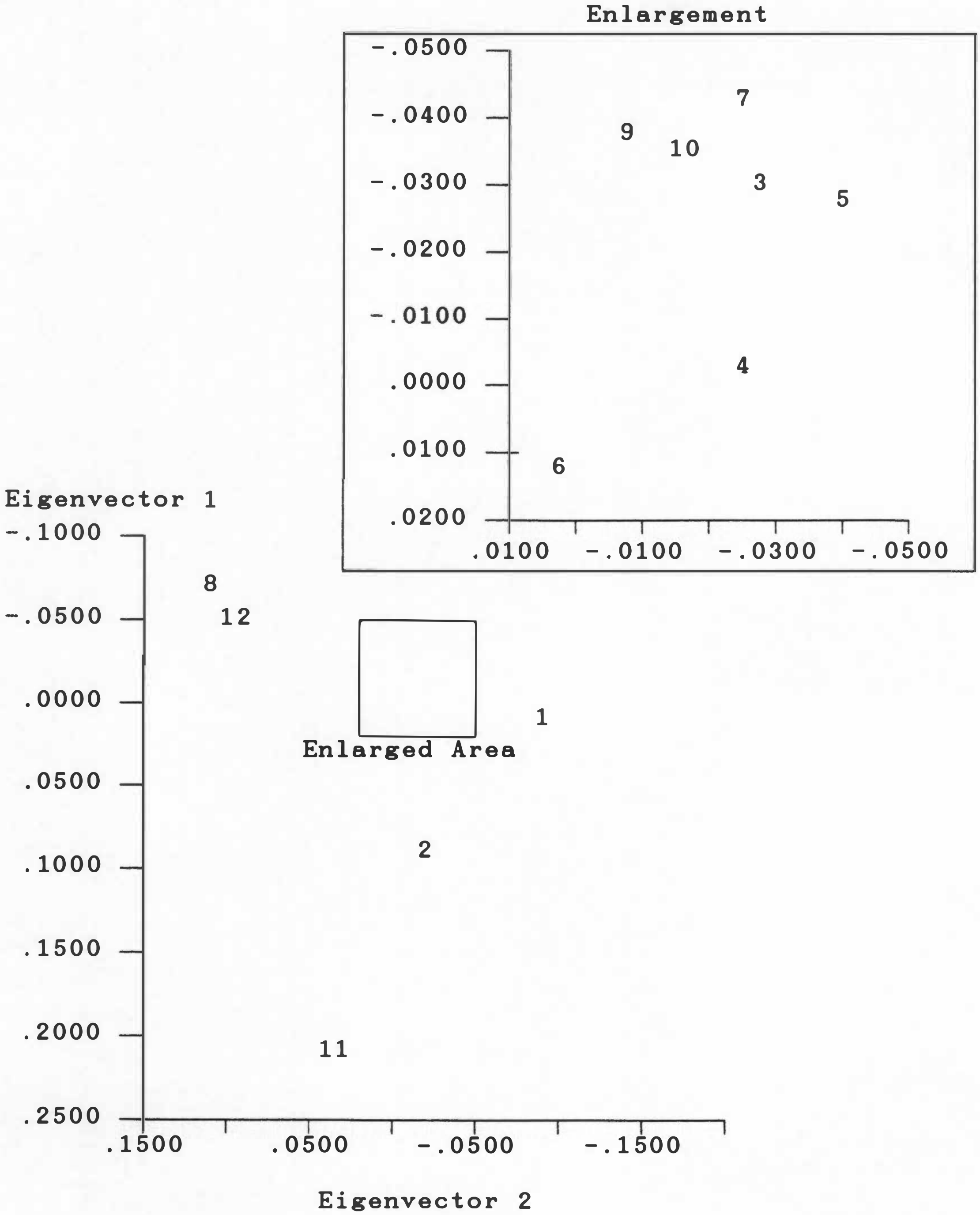
Average d^2 Distances and Rankings for Each District
Sevier County, Tennessee

District	1860 Avg.	1860 Rank	1870 Avg.	1870 Rank	1880 Avg.	1880 Rank	Mean Rank
1. Jones Cove	.042	3	.051	4	.049	2	3
2. Emerts Cove	.042	3	.071	2	.041	3	3
3. Fair Garden	.027	11	.037	9	.033	5	8
4. Harrisburg	.024	12	.033	13	.023	15	13
5. Sevierville	.030	8	.030	14	.023	15	12
6. Wears Valley	.029	9	.057	3	.027	11	8
7. Catlettsburg	.031	7	.035	11	.029	9	9
8. Henrys Xroads	.039	4	.039	8	.031	8	7
9. Trundles Xroads	.028	10	.034	12	.025	13	12
10. Cusicks Xroads	.032	6	.036	10	.026	12	9
11. Gatlinburg	.068	1	.087	1	.073	1	1
12. Sinking Spgs.	.036	5	.033	5	.033	5	5
13. Richardsons C.			.043	7	.032	6	7
14. Boyds Creek			.043	7	.028	10	8
15. Allensville					.031	8	8

relationships among groups. This has been done the census years 1860, 1870, and 1880.

The graph or "surname map" for 1860 is presented as Figure 3 on the following page and coordinates for each district for all three censuses are presented in Table XIII.

Two eigenvectors account for almost half of all variation in 1860. The first, presented on the vertical axis, summarizes 31.07 percent of total variation and separates Gatlinburg from all other districts. Emerts Cove is also separated by this variate, although its position is closer to the other districts than to Gatlinburg. The second



1. Jones Cove	5. Sevierville	9. Trundles Xroads
2. Emerts Cove	6. Wears Valley	10. Cusicks Xroads
3. Fair Garden	7. Catlettsburg	11. Gatlinburg
4. Harrisburg	8. Henrys Xroads	12. Sinking Springs

Figure 3. Plot of Principal Coordinates, Sevier County Civil Districts: 1860

TABLE XIII

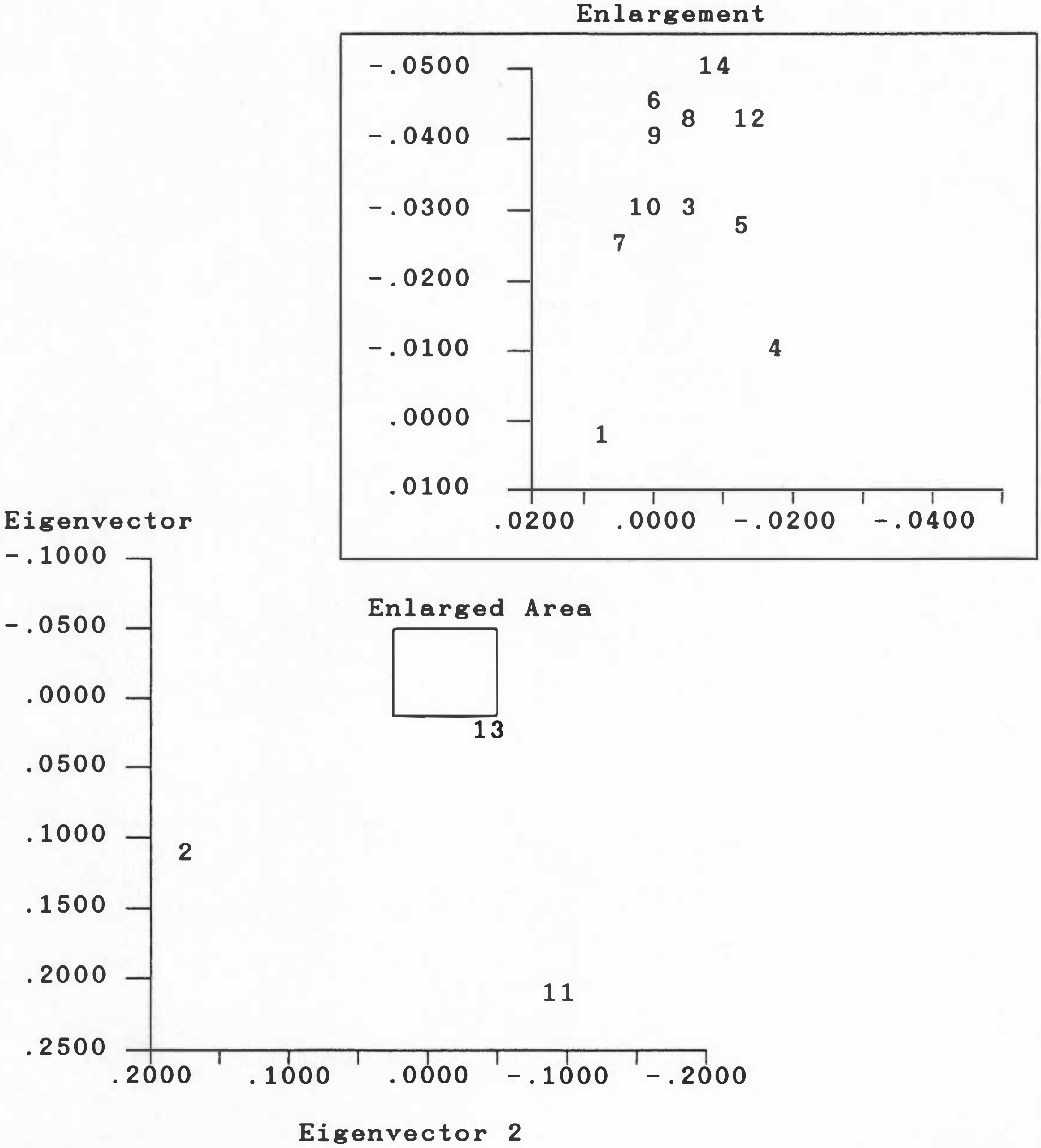
Principal Coordinates Analysis for Each District, Sevier County, Tennessee: 1860, 1870, 1880

Year	1860		1870		1880	
Eigenvector	1	2	1	2	1	2
District						
1. Jones Cove	-.0040	-.0803	.0006	.0082	.0004	-.1639
2. Emerts Cove	.0835	-.0203	.1062	.1816	-.0559	-.0280
3. Fair Garden	-.0304	-.0285	-.0307	-.0043	.0350	.0095
4. Harrisburg	-.0041	-.0246	-.0108	-.0188	.0045	.0087
5. Sevierville	-.0270	-.0410	-.0297	-.0119	.0171	-.0026
6. Wears Valley	.0127	.0025	-.0460	-.0017	.0216	.0031
7. Catlettsburg	-.0436	-.0250	-.0266	.0051	.0407	.0377
8. Hnys. Xroads	-.0609	.1041	-.0425	-.0055	.0411	.0120
9. Tdls. Xroads	-.0375	-.0082	-.0409	.0000	.0324	.0176
10.Csks. Xroads	-.0352	-.0142	-.0314	.0024	.0244	.0097
11.Gatlinburg	.2021	.0436	.2214	-.0904	-.2186	.0443
12.Snkng. Spgs.	-.0557	.0919	-.0428	-.0131	.0261	.0199
13.Rchrdsns. C.			.0221	-.0444	-.0517	-.0335
14.Boyds Creek			-.0490	-.0071	.0417	.0220
15.Allensville					.0412	.0436

eigenvector reflects 16.07 percent of variation and divides Henrys Crossroads and Sinking Springs from all other districts.

On the graph (Figure 3) districts in the center of the county--Fair Garden, Sevierville, Catlettsburg, Trundle's Crossroads, and Cusicks Crossroads--form one cluster (see enlargement), while districts 8 and 12 form another. The mountain districts do not cluster but are scattered across the graph. Spatial segregation from other districts is pronounced for Emerts Cove and is extreme for Gatlinburg.

Gatlinburg's spatial isolation is even more evident in the problematic 1870 graph (Figure 4), where it is an outlier



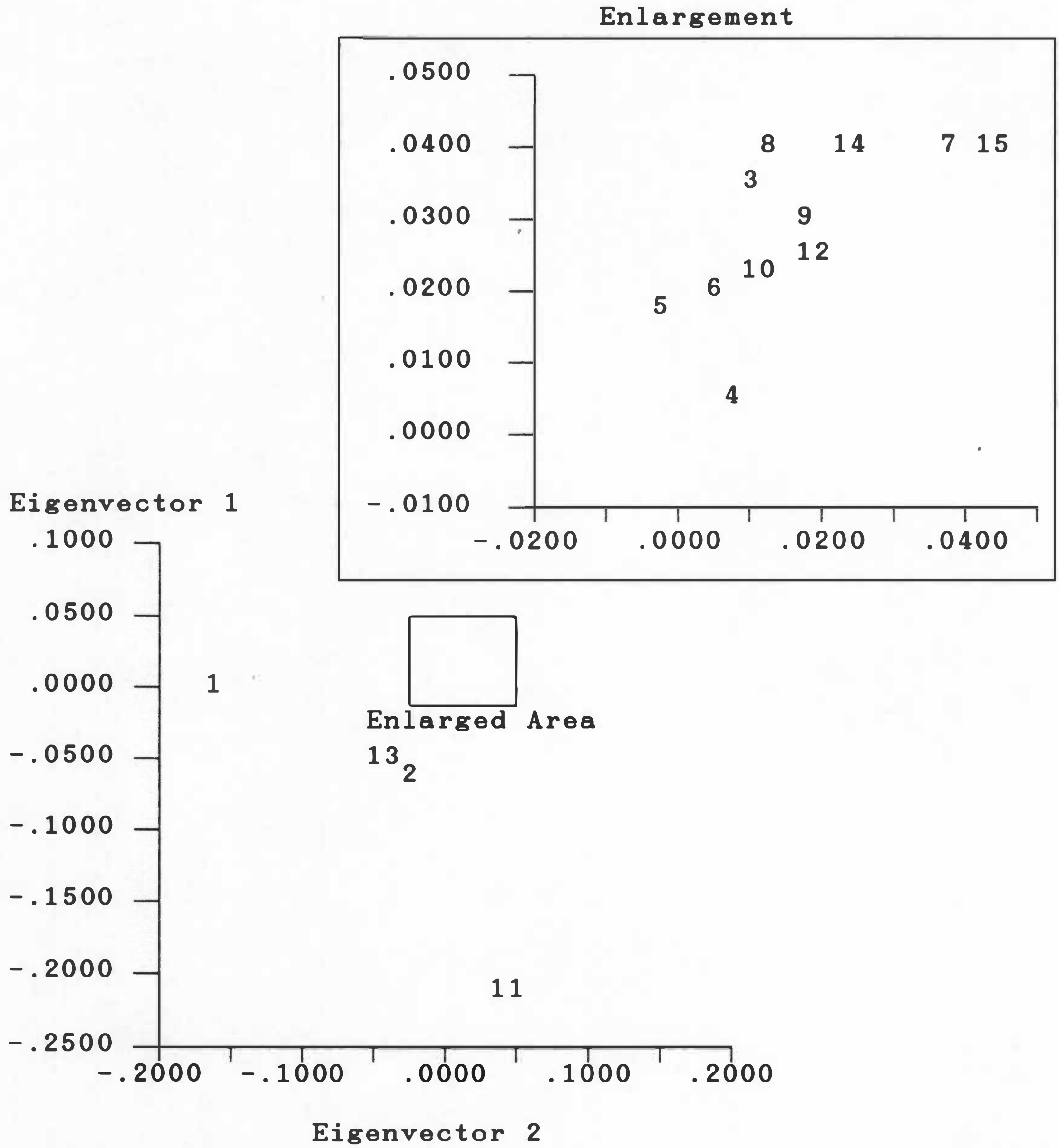
1. Jones Cove	6. Wears Valley	11. Gatlinburg
2. Emerts Cove	7. Catlettsburg	12. Sinking Springs
3. Fair Garden	8. Henrys Xroads	13. Richardsons Cove
4. Harrisburg	9. Trundles Xroads	14. Boyds Creek
5. Sevierville	10. Cusicks Xroads	

Figure 4. Plot of Principal Coordinates, Sevier County Civil Districts: 1870

on both orthogonal variates. Emerts Cove lies midway between Gatlinburg and a cluster of districts on eigenvector 1. On eigenvector 2, however, Emerts Cove is at the opposite extreme of Gatlinburg, while the main cluster of districts lies near the centroid. This cluster consists of Fair Garden, Sevierville, Wears Valley, Catlettsburg, Trundles Crossroads, Cusicks Crossroads, and Boyds Creek. The districts north of the French Broad, which formed a separate group in the 1860 graph, are still closely associated, but are included within the main array. Jones Cove, Harrisburg, and the new district of Richardsons Cove are not included in this cluster, but are proximate. Eigenvectors 1 and 2 account for 25.01 percent and 14.81 percent of the total variation, respectively.

Noteworthy are the positions of the newly created districts relative to their parent districts. Boyds Creek lies close to Trundles Crossroads on eigenvector 1, while Richardsons Cove lies between its progenitors, Emerts Cove and Harrisburg (although closer to the latter), on the same axis.

Coordinates for the two major variates in 1880 are plotted in Figure 5 and represent over 40 percent of all variation--28 percent by eigenvector 1 and 15 percent by eigenvector 2.



1. Jones Cove	6. Wears Valley	11. Gatlinburg
2. Emerts Cove	7. Catlettsburg	12. Sinking Springs
3. Fair Garden	8. Henrys Xroads	13. Richardsons Cove
4. Harrisburg	9. Trundles Xroads	14. Boyds Creek
5. Sevierville	10. Cusicks Xroads	15. Allensville

Figure 5. Plot of Principal Coordinates, Sevier County Civil Districts: 1880

Gatlinburg is again a gross outlier on the vertical axis. A preponderance of districts associate at the other extremity of eigenvector 1, with the near convergence of Henrys Crossroads, Catlettsburg, Boyds Creek and Allensville. The relationship between Emerts and Richardsons Coves is evident as they cluster on both eigenvectors.

The outlier on eigenvector 2 is Jones Cove while Gatlinburg lays between Catlettsburg and Allensville at the opposite side of the graph. Districts 3, 4, 5, 6, 8, 9, 10, 12, and 14 are positioned closely on this eigenvector.

Interpreting these genetic or surname maps is difficult. Because of the exclusion of less significant variates, some information is not included. Additionally, district samples are small and unequal. Finally, the possibility of systematic error in the 1870 data makes particularly suspect any relationship observed solely in that map.

Thus it is difficult to say with any certainty what a particular configuration means. Does the close association in 1860 of districts north of the French Broad signify that the river acted as a genetic barrier before the Civil War? Is the close association in 1860 of Boyds Creek, Henrys Crossroads, Allensville and Catlettsburg on eigenvector 1 reflective of the formation of a riverine community or gene pool? It is hard to discern whether these patterns are real or spurious.

But one element is as apparent in the surname/genetic maps as it was in the simple genetic distances and in the measures of within-group isonomy. The isolation of the mountainous districts is conspicuous in all years. Jones Cove and Emerts Cove show a high level of internal homogeneity with respect to surnames, but Gatlinburg is by far the most homogeneous in that regards. If the assumption of a surname-genetic analogy holds true, Gatlinburg was the most genetically isolated as well.

CHAPTER 7

Discussion and Conclusions

The Whaley family came to Emerts Cove from North Carolina early in the 19th century. A religious folk, they faithfully followed the Word of God, particularly Genesis 1, verse 28 which states "Go, be fruitful and multiply." The family was so true to this charge that soon the valley was filled with Whaleys. Their descendants, now scattered across the nation, still refer to themselves as "the Whaleys of Emerts Cove[1]".

A story is still told in the cove of a traveling salesman who, after an arduous day peddling goods to the poor cove farmers, came upon the log cabin of a man named Bill Guess. The salesman walked up to the porch where Bill was sitting, introduced himself and asked Bill his name. "Guess" came the reply. Without a moment's hesitation, the salesman responded "Well, I'd guess it's Whaley"[2].

The Whaleys were not the only large family in Emerts Cove; There were also the Huskeys and the Shultzes. In the hills around Gatlinburg lived more Whaleys, along with Ogles, McCarters, and Kings, while Jones Cove was filled with Hursts,

---

[1]The information about the Whaley Family was taken from a newspaper clipping entitled "The Whaleys of Greenbriar" by Carson Brewer, a staff writer at the Knoxville News-Sentinel. The date of publication is not known as the clipping was provided by a member of the Whaley and is not dated.

[2]Ibid.

Breedens, and Williamses. It was the presence of these extended families, whose members lived in close proximity, that led early writers on Appalachia to talk of mountain clans, intermarriage, and inbreeding.

> Owing to isolation of the clans, and their extremely limited travels, there are abundant cases like those caustically mentioned in *King Spruce*: "All Skeets and Bushees, and married back and forth and crossways and upside down till every man is his own grandmother if he only knew enough to figger relationship" (Kephart 1976:297).

The existence of such large families in the mountain districts is reflected in the high rates of isonomy, high kinship coefficients, and genetic isolation recorded by the present study. What caused this clustering of kinfolk and homogeneity of surnames?

Evidence from the marriage records suggests that breeding subdivision may have played a part. The high ratio of nonrandom to random components in the inbreeding coefficients and in the repeated pairs analysis implies "a tendency to limitations in the choice of partners, a reduction in genetic variability...(Lasker et al. 1986:421)".

Isolation by distance--an inverse correlation of the distance separating communities or individuals and the probability of marriage between those communities or individuals--would result in some departure from panmixia and probably contributed to the nonrandom nature of mate selection in Sevier County. The higher mean kinship in the mountainous districts, however, demonstrates that factors more complex than simple distance isolation were operating. The mountains were secluded--the introduction of new families and new surnames, (if not new genes) was inhibited by barriers. Exactly what form did these barriers take?

One possible explanation for the mountaineers' genetic isolation could be the cultural and economic differences that divided lowland districts from the mountains. Lowland farmers had more acreage under cultivation, produced a large surplus crop, and could easily move their crops to market. The hill farmer could not produce a substantial surplus; the bulk of his produce going to feed his own family. Mountaineers supplemented their incomes and food supply by hunting, gathering plants, distilling corn whiskey, and after 1900, by logging. An economic dichotomy between lowland farmers and mountaineers was noted by Ales Hrdlicka in his book Old Americans.

> The third group is that of the Tennessee mountaineers with some of the lower land people. The latter are in the main farmers; the former are partly farmers, partly hunters, partly laborers (1925:184).

Similar distinctions in subsistence and culture have been noted between ridge dwellers and valley farmers in other Appalachian communities (Matthews 1965).

Such cultural differences could have resulted in breeding segregation between groups. In Yorkshire, England, one study of isonomy revealed a sharp division between family names common among sailors, shipowners and fishermen, and those surnames common among farmers, laborers, and alum workers (Smith and Hudson 1984). Marriage and gene flow were constrained by occupational orientation. The distinctions between highland and lowland communities in Sevier County were not as great as between Yorkshire's mariners and landlubbers, but the differences that did exist may have had a like, if less intense, effect on gene flow.

Terrain must have played some role in this genetic isolation. Conventional wisdom has been that social relationships in the mountains tended to be localized because of the difficulty of travel. Anecdotal evidence, however, suggests that wider alliances and travels were not unknown. This is illustrated by the tale of one Gatlinburg resident who, as a youngster in the 1890's, ran away from home.

> When my father arrived home in the mountains first thing wanted to know about his son Wiley and my people told him I had disapered (*sic*) one night and no dought (*sic*) I had went to N.C. to some of our mothers people or I had went to K.Y. where I had an uncle....but when I was located I were among the Indians in N.C. eating bear meat and other wild animals meat ...(Oakley 1940:42-43)

Transportation hardships apparently did not prevent "courting" and marriage between members of isolated mountain communities. An earlier study of intermarriage in the Gatlinburg district documented such marital exchanges.

> Because of the mountains, the roads were made following the rivers, and not infrequently the beds of shallow streams served as roads. This meant that during heavy rains the roads were practically useless, and these valleys were cut off from the rest of the county for weeks at a time. It also restricted intercourse between the two valleys and visiting between the inhabitants of the two was rare, although the Little River Valley was settled by people from the Little Pigeon Valley and intermarriage between the two was constantly taking place (Carter 1928:458).

That the inhabitants of these two communities were intermarrying suggests that proximity was not the sole determinant of mate selection. The Little Pigeon Valley (Gatlinburg) and the Little River Valley (now called Elkmont) were less than 10 miles apart overland, but because of the terrain, the distance by road was 21 miles (Carter 1928:458).

The ancestral connection of these two communities evidences a lineage-based preference in mate selection and supports the notion of clannish behavior. Some writers have claimed that the Southern Mountaineer had a preference for marriage with those of "his own bloodline" (Caudill 1962:84), because years of feuding had led to animosity between families. Other studies have uncovered endogamous mating patterns, but offer less simplistic explanations for this "familism".

Several ethnologists have shown how kinship in Appalachia affected social relationships and residence patterns within the local community and also between communities (Pearsall 1959; Matthews 1965; Bryant 1981). Mate choice, church or work-group membership, even land tenure are features of Appalachian culture often governed by kinship. Although their importance varies from community to community, kinship and the kin-group are fundamental structural elements of Appalachian society[3].

The often mentioned independence of the Southern Mountaineer would be better described as interdependence. Farming, clearing land, construction, and politicking were activities that required group cooperation, especially in "cash-poor" 19th and early 20th century Appalachia. Kin-groups formed basic social units around which such activities could be organized, and membership in a kin-group involved both obligations and expectations with respect to such activities (Matthews 1965; Bryant 1981; Hicks 1976:36-38).

Relationship was a major element in the formation of kin-sets, but this formation was somewhat arbitrary. One second cousin might be recognized as kin while another might not (Batteau 1982). Kinship channeled individuals into sets and

[3]This statement and others following relate cultural features common in Southern Appalachia. They are not meant as descriptors of a universal culture pattern. No simplistic statement can adequately portray a region as extensive and diverse as Appalachia.

facilitated the development of such groups, but was not a final determinant of membership; dwelling proximity and personality were important as well.

Sometimes certain "moral and reputational features" became associated with a particular surname (Batteau 1982:455). A family's "reputation," if positive, had benefits for individuals included within the kin-set.

> "He came from that set of Lewises on Rabbit Creek. I knowed he'd be good for the money, and I let him put his cabbage on five acres of mine (Hicks 1976:39)

Conversely a set might be viewed as "no-count"; dishonest, fractious, or just generally disagreeable. In such cases the consequences for individuals might include a limitation of potential mates. Folk from "good" families would be hesitant to marry folk from undesirable ones.

In this case, surnames would not be selectively neutral. Individuals with high-status "good" surnames would have a breeding advantage over individuals with "bad" surnames. The limits in mate choice resulting from a bad "moral reputation" may have forced some lower status families in Appalachia to become more endogamous--more inbred. This has been documented for at least one "unprestigious" set (Batteau 1982:451)". Such limitations, combined with geographical seclusion, probably induced the development of isolate populations.

Many of the early reports of inbreeding in Appalachia describe specific "clans" or kin-sets (Manne 1936)--perhaps the Melungeons originated in this manner. Certainly status segregation could explain the high proportion of repeated pairs in marriages found in certain Sevier County lineages (Table VII).

Kinship, combined with economics, may have encouraged endogamy in other ways. The economic hardships that restricted immigration into the mountains would have been lessened for individuals with strong kin-group affiliations. In Sevier County the average size of farms decreased from 270 acres in 1860 to 78 acres in 1900, while the average number of acres that were "improved" decreased from 66 to 36 (Cummings 1988:37). Because land was usually handed down from parent to child, the choice of a mate from within an individual's lineage would have kept family land under the lineage's control. Affiliations and mate exchange between lineages might have insured an adequate supply and workable distribution of land (Matthews 1965). A preference for kin or clan in marriage and residence, if it indeed existed, would have had profound effects on both the genetic structure of the population and the distribution of surnames.

The land shortage may also have led individuals who married outsiders to leave the mountains for residence in their mates' communities. Individuals with weak kin-ties to a mountain community may have left for the greater social and economic opportunities existing in cities or in the opening American West. Such a phenomenon apparently took place in Cades Cove in neighboring Blount County where the number of tenant farmers decreased after 1870 as did the number of different surnames (Dunn 1988:70-73). If selective emigration from Sevier County took place, then the increased homogeneity of surnames may reflect decreased variation in the gene pool resulting from the departure of individuals with rarer alleles.

But a full investigation of these possibilities is beyond the scope of the present study. Surname isonomy alone does not readily explain the reasons for an increase in genetic homozygosity or even prove its existence--it is only an indicator of that homozygosity. Nonrandom isonomy may reveal the degree of mating subdivision in a population, but it does not clearly delineate the nature of that subdivision.

Also, as previously detailed, there are problems with these measures. Polyphyletic surnames and virilocal residence patterns can inflate inbreeding estimates calculated from isonomy. Studies comparing pedigree-derived coefficients with coefficients derived from isonomy have found that the latter frequently overestimate inbreeding and kinship (Friedl and Ellis 1974; Hussels 1969; Jorde and Morgan 1987). Since no cohesive, comprehensive, body of genealogical data exists for Sevier County, it is impossible to test the accuracy of the kinship and inbreeding coefficients county-wide.

For the Gatlinburg district alone, however, comparative data do exist. A 1924 study by Isabel Gorden Carter found "considerable inbreeding" in the most mountainous part of Sevier County. Carter examined genealogies of 302 11th District residents from four to twenty years in age and was able to identify 92 "distinct family strains" in the population.

> These ninety-two ancestors are not related to one another except by marriage or through their descendants...They appear as ancestors and are scattered from the second to the ninth generation. Four of them have contributed to over 90 per cent of the total number of fraternities studied, fifteen are common to over 30 percent of them, and forty appear in less than 3 per cent of the entire number. (1928:465)

Carter also determined the number of ancestors for each individual for six generations, and then averaged these estimates. Table XIV displays these averages for six generations along with the theoretical number of ancestors for an individual from a lineage with no intermarriage. Also displayed is the number of ancestors an individual would have if he were the offspring of a first cousin marriage.

Determining the exact amount of inbreeding from the average number of ancestors in a generation is problematic because one individual may appear in different generations of a single genealogy (Carter 1928:462). Still, the parents of

TABLE XIV

Average and Theoretical Number of Ancestors for Six Generations

	# of Ancestors, No Inbreeding	# of Ancestors, Child of 1st Cousins	# of Ancestors for Gatlinburg Youth, 1928[a]
Generation			
2nd	4	4	3.97
3rd	8	6	7.78
4th	16	12	13.87
5th	32	24	24.26
6th	64	48	46.96

[a]Derived from the table on page 465 of Carter's study.

Gatlinburg children born from 1904 to 1920[4] appear to have had cumulative kinship approaching that of first cousins.

If no kinship preference governed mating within the community, then mean *a priori* kinship for the parent's generation would approximate .0625, the kinship of first cousins. This is higher than the level of kinship indicated by the surname estimates which ranged from .0151 to .0183.

It is unlikely that an increase in mean kinship from the 1880 census to the early 20th century can fully account for this discrepancy. Carter's method of counting individuals who appeared in different generations of the same genealogy might

---

[4]This range was determined by subtracting the age spread of the study subjects, 4 to 20 years, from 1924, the year of the study.

have resulted in a slight underestimation of the number of ancestors[5], or sampling error in the surname estimates could be responsible for some of the disparity. A more probable explanation of the greater pedigree estimates, however, is lineage endogamy.

Mean marital kinship that is greater than mean district kinship indicates that mating patterns were subdivided within the district. Lineage or kin preference in mate selection results in higher mean kinship between mates (the pedigree estimate) than between all residents of the district (the surname estimate). Apparently the mountain inhabitants of the 11th district were practicing lineage or kin-set endogamy to some degree.

Such preference would have raised individual homozygosity in the local population. What physical effects would be expected? In theory genotypic variation should decrease, and genetic diseases arising from recessive alleles might increase. In actuality, such effects were not clearly evident.

Carter documented low variation of the cephalic index in Gatlinburg families compared with variation reported for families in other populations (1928:465-468). Her study measured children, but the cephalic index may change slightly from childhood to adulthood. If age is a source of variation

---

[5]Carter counted such an individual only in the most recent generation in which he/she appeared.

for this index, then Carter's comparison of coefficients derived from children with coefficients derived from adults may not be valid.

As far as health, fitness, and mental capacity, we must rely on anecdotal information provided by Carter.

> Dr. J.W. Ogle, the physician who is most familiar with the area...gives it as his opinion that the general health is slightly below the average elsewhere and that the most common ailments are the ordinary respiratory diseases (colds, catarrh, and pneumonia) and gasto-intestinal disorders. In his opinion, the mentality of the inhabitants is above the average and the percentage of feeble-mindedness less than elsewhere. Children of 1st cousin marriages do not, in his opinion, measure up to community standard, although they are stronger physically and mentally than he had expected to find them...Miss Evelyn Bishop states that she concurs with Doctor Ogle as regards the mentality and feeble-mindedness in this district. Miss Bishop has been in charge of the Pi Beta Phi Settlement School for some years. (Carter 1928:459-460).

Apparently the hills were not filled with the mental idiots, illness and "the worst in sub-representative types of humanity" that according to some researchers infested other parts of Appalachia (Hirsch 1928:237).

Breeding subdivision sorts alleles into homozygotic pairs rather efficiently and along with isolation may have created pockets of genetic disease and homozygosity in some parts of Appalachia. Clan endogamy in particular has resulted in particular lineages with elevated levels of inbreeding.

Still, the elimination of alleles depends on the action of selection or drift[8]. A lack of time depth since the settlement of Sevier County precluded these processes from having any tremendous effect on the gene pool. There was no "inbreeding depression."

Gatlinburg, for example, had been settled for no more than six generations at the turn of the century, and its isolation was disappearing even at the time of Carter's study.

> Since 1924 remarkable changes have taken place in the immediate vicinity of Gatlinburg. Excellent roads with the prosperity growing out of large numbers of summer visitors have facilitated a rapid cultural modification. Electric lights, painted homes, Ford cars, greater variety in dress, an occasional movie and frequent trips out of the mountains indicate what is taking place (Carter 1928:459).

This was true of the rest of Appalachia as well. Like tourism in Sevier County, the timber and coal industries brought new economic opportunities and new immigrants to other parts of the region. Roads and automobiles ended the physical isolation, just as radios and telephones ended cultural isolation. Any increase in homozygosity brought about by genetic isolation and population subdivision has certainly been erased by the subsequent social and cultural changes that swept the region.

---

[8]Differential migration, as previously discussed, may be considered a kind of selection.

Surname distributions can shed light on many aspects of a population's genetic structure. In Sevier County surnames have exposed population subdivision, micro-differentiation, and variation in average kinship from district to district. The level of isonomy and genetic isolation observed among the highland dwellers indicates the effect of geography on the genetic structure of the county. The mountains did indeed restrict gene flow. The frequency of repeated pairs in marriages (especially in certain lineages) and the discrepancy between estimates of kinship from surnames and estimates from genealogies indicate that kin preference in mate selection may have played some role in any increased homozygosity. In the mountains of Sevier County genetic structure mirrored social structure.

Sevier County, however, is only one small section of Appalachia. Research in other communities has produced results and conclusions that differ markedly from those of this study. It may be that Appalachia is too diverse geographically, politically, and historically for any one region to be considered truly representative.

The present study only reveals the genetic structure of Sevier County. While the results may prove useful for theory building and comparison, and may provide a framework for future research, only a series of such studies can provide a full understanding of population genetics in Southern Appalachia.

LIST OF REFERENCES

List of References

Allen, G.
  1988 Random Genetic Drift Inferred from Surnames in Old Colony Mennonites. Human Biology 60: 639-653.

Azevado, E., N. E. Morton, C. Miki, S. Yee
  1969 Distance and Kinship in Northeastern Brazil. American Journal of Human Genetics 21:1-22.

Batteau, A.
  1982 Mosbys and Broomsedge: the Semantics of Class in an Appalachian Kinship System. American Ethnologist 9:445-446.

Brues, A.M.
  1977 People and Races. Macmillian Publishing Inc. New York.

Bryant, C.F.
  1981 We're All Kin: A Cultural Study of a Mountain Neighborhood. University of Tennessee Press. Knoxville.

Carter, I.G.
  1928 Reduction of Variability in an Inbred Population. AJPA 11:457-471 (old series).

Caudill, H.M.
  1962 Night Comes to the Cumberlands: A Biography of a Depressed Area. Little, Brown, and Co. Boston.

  1976 A Darkness at Dawn: Appalachian Kentucky and the Future. University of Kentucky Press. Lexington.

Chakraborty, R.
  1985 A Note on the Calculation of Random RP and Its Sampling Variance. Human Biology 57:713-717.

Chen, Kuang-Ho and L.L. Cavali-Sforza
  1983 Surnames in Taiwan: Interpretations Based on Geography and History. Human Biology 55:367-374.

Crow, J.F.
  1983 Discussion. In: Surnames as Markers of Inbreeding and Migration. (arranged by K. Gottlieb). Human Biology 55:383-397.

Crow, J.F., and A.P. Mange
1965 Measurement of Inbreeding from the Frequency of Marriages between Persons of the Same Surname. Eugenics Quarterly 12:199-203.

Cummings, J.
1988 Community Violence and the Nature of Change: White Capping in Sevier County, Tennessee in the 1890's. Masters Thesis, University of Tennessee. Knoxville.

Devor, E.J.
1980 Marital Structure and Genetic Isolation in a Rural Hispanic Population in Northern New Mexico. AJPA 53:257-265.

Devor, E.J., M.H. Crawford, and T. Koertvelyessy
1983 Marital Structure and Genetic Heterogeneity of Ramera Island, Newfoundland. AJPA 61:401-419.

Dunn, D.
1983 The Folk Culture of Cades Cove, Tennessee. Tennessee Folklore Society Bulletin 43(2):67-87.

1988 Cades Cove, The Life and Death of a Southern Appalachian Community: 1818-1937. University of Tennessee Press. Knoxville.

Dyke, B., A.V. James, and W.T. Morrell
1983 Estimation of Random Isonomy. Annals of Human Biology 10:295-298.

Friedl, J., and W.S. Ellis
1974 Inbreeding, Isonomy, and Isolation in a Swiss Community. Human Biology 46:699-712.

Fuster, V.
1986 Relationship by Isonomy and Migration Pattern in Northwest Spain. Human Biology 58:391-406.

Gower, J.C.
1972 Measures of Taxonomic Distance and Their Analysis. In: J.S. Weiner and J. Huizinga, (eds.). The Assessment of Population Affinities in Man. pp. 1-24. Clarendon Press. Oxford.

Hicks, G.L.
1976 Appalachian Valley. Holt, Rinehart, and Wilson. New York.

Hirsch, N.D.
1928 An Experimental Study of the East Kentucky Mountaineers. Genetic Psychology Monographs 3(3):183-244.

Hrdlicka, A.
1925 The Old Americans: A Physiological Profile. Reprint: Arno Press and the New York Times, 1970. New York.

Hulse, F.S.
1958 Selection for Skin Color among the Japanese. AJPA 27:145-159.

Hurd, J.P.
1983 Comparison of Isonomy and Pedigree Analysis Measures in Estimating Relationships between Three 'Nebraska' Amish Churches in Central Pennsylvania. Human Biology 55:349-355.

Hussels, I.
1969 Genetic Structure of Saas, a Swiss Isolate. Human Biology 41:469-479.

James, A.V., G.W. Lasker and W.T. Morrell
1986 A Test of the RP Method on St. Bart, F.W.I.. (Abstract). AJPA 69:219.

Jorde, L.B., and K. Morgan
1987 Genetic Structure of the Utah Mormons: Isonomy Analysis. AJPA 72:403-412.

Kamizaka, M.
1954 Frequency of Isonomous Marriages. Seibutsu Tokei-gaku 2:292-298 (in Japanese).

Kephart, H.
1976 Our Southern Highlanders: A Narrative of Adventure in the Southern Appalachians and a Study of Life among the Mountaineers. University of Tennessee Press. Knoxville. (originally published in 1913).

Kirkland, J.R., and R.L. Jantz
1977 Inbreeding, Marital Movement, and Genetic Isolation of a Rural Appalachian Population. Annals of Human Biology 4:211-218.

Koertvelessy, T., M.H. Crawford, R.G. Huntsman, M. Collins, D. Kelping, and M. Uttley
1988 Repetition of the Same Pair of Surnames in Marriages in Fogo Island, Newfoundland, and Genetic Variation. AJPA 77:253-260.

Lasker, G.W.
1985 Surnames and Genetic Structure. Cambridge Studies in Biological Anthropology, 1. Cambridge University Press. Cambridge.

Lasker, G.W., and B.A. Kaplan
1985 Surnames and Genetic Structure: Repetition of the Same Pairs of Names of Married Couples: A Measure of Subdivision of the Population. Human Biology, 57:431-440.

Lasker, G.W., C.G.N. Mascie-Taylor and D.A. Coleman
1986 Surnames and Genetic Structure: Repetition of the Same Pair of Names of Married Couples, a Measure of Subdivision of the Population. Human Biology 57:713-717.

Lasker, G.W.
1988 Repeated Surnames in Those Marrying into British One-Surname "Lineages": An Approach to the Evaluation of Population Structure through Analysis of the Surnames in Marriage. Human Biology 60:1-9.

Libet, M.L.
1983 Expected Versus Actual Random Inbreeding: A Reinterpretation of the Random/Nonrandom Inbreeding Formula. Human Biology 55:323-329.

Lower, M.A.
1988 A Dictionary of Surnames. Wordsworth Editions Ltd. Hertfordshire, England. (First published as *A Dictionary of the Family Names of the United Kingdom* by John Russell Smith. G.P. Bacon, Lewis. London.

Manne, J.
1936 Mental Deficiency in a Closely Inbred Mountain Clan. Mental Hygiene 20:269-279.

Mascie-Taylor, C.G.N., G.W. Lasker and A.J. Boyce
1987 Repetition of the Same Surnames in Different Marriages as an Indication of the Structure of the Population of Sanday Island, Orkney Islands. Human Biology 59:97-102.

Matthews, E.M.
1965 Neighbor and Kin. Life in a Tennessee Ridge Community. Vanderbilt University Press. Nashville.

McCullough, J.M., E. Giles and R. Thompson
1985 Evidence for Assortive Mating and Selection in Surnames: A Case from Yucatan, Mexico. Human Biology 57:375-386.

Oakley, W.
1940 Roamin' & Restin' with the Roamin' Man of the Smoky Mountains. Oakley Enterprises, Inc. Sevierville.

Pearsall, M.
1959 Little Smoky Ridge: The Natural History of a Southern Appalachian Neighborhood. University of Alabama Press. Birmingham.

Pollitzer, W.S., and W.H. Brown
1969 Survey of Demography, Anthropometry, and Genetics in the Melungeons of Tennessee: An Isolate of Hybrid Origin in the Process of Dissolution. Human Biology 41:388-399.

Relethford, J.H.
1988 Estimation of Kinship and Genetic Distance from Surnames. Human Biology. 60:475-492.

Schreider, E.
1969 Inbreeding, Biological and Mental Variations in France. AJPA 30:215-220.

Smith, M.T., and B.L. Hudson
1984 Isonymic Relationships in the Parish of Fylingdales, North Yorkshore, in 1851. Annals of Human Biology 11(2):141-148.

Smoky Mountain Historical Society
1984 In the Shadow of the Smokies: Sevier County, Tennessee Cemeteries. Smoky Mountain Historical Society. Sevierville.

1986 The Gentle Winds of Change: a History of Sevier County, Tennessee, 1900-1930. Smoky Mountain Historical Society. Sevierville.

Sorg, M.H.
1983 Isonomy and Diabetes Prevalence in the Island Population of Vinalhaven, Maine. Human Biology 55:305-311.

Swedlund, A.C., and A.J. Boyce
1983 Mating Structure in Historical Populations: Estimation by Analysis of Surnames. Human Biology 55:251-262.

Tincher, R.B.
1980 Night comes to the Chromosomes: Inbreeding and Population Genetics in Southern Appalachia. Central Issues in Anthropology 7:27-50.

Trout, E.
1984 Gatlinburg: Cinderella City. Griffen Graphics. Maryville, Tennessee.

Underwood, J.H.
1979 Human Variation and Human Microevolution. Prentice-Hall Inc. Englewood Cliffs, California.

Weis, K.H., R. Chakraborty, A.V. Buchanan, R. Schartz
1983 Mutations in Names: Implications for Assessing Identity by Descent from Historical Records. Human Biology 55:313-322.

Weyl, N.
1974 The Geography of Stupidity in the U.S.A. Mankind Quarterly 15:117-123.

Wilson, S.T.
1914 The Southern Mountaineers. Presbyterian Home Missions Literature Department. New York.

Yasuda, N.
1983 Studies of Isonomy and Inbreeding in Japan. Human Biology 55:330-338.

APPENDIX

TABLE A-I

Distance (d^2) Matrix of Sevier County, Tennessee: 1860

Districts	1	2	3	4	5	6	7	8	9	10	11	12
1 (N = 171)	.000	.042	.030	.029	.033	.040	.040	.048	.035	.042	.077	.047
2 (N = 101)	.042	.000	.036	.033	.039	.038	.044	.054	.039	.044	.048	.047
3 (N = 228)	.030	.036	.000	.012	.018	.019	.017	.031	.018	.019	.068	.028
4 (N = 216)	.029	.033	.012	.000	.012	.016	.015	.031	.017	.021	.054	.028
5 (N = 186)	.033	.039	.018	.012	.000	.021	.022	.037	.019	.024	.070	.032
6 (N = 146)	.040	.038	.019	.016	.021	.000	.024	.035	.022	.024	.051	.031
7 (N = 179)	.040	.044	.017	.015	.022	.024	.000	.034	.019	.022	.075	.033
8 (N = 140)	.048	.054	.031	.031	.037	.035	.034	.000	.030	.037	.082	.015
9 (N = 192)	.035	.039	.018	.017	.019	.022	.019	.030	.000	.014	.069	.025
10 (N = 154)	.042	.044	.019	.021	.024	.024	.022	.037	.014	.000	.071	.031
11 (N = 95)	.077	.048	.068	.054	.070	.051	.075	.082	.069	.071	.000	.079
12 (N = 86)	.047	.047	.028	.028	.032	.031	.033	.015	.025	.031	.079	.000

Table A-II

Distance (d^2) Matrix of Sevier County, Tennessee: 1870

Districts	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1 (N = 165)	.000	.073	.044	.038	.037	.066	.044	.044	.044	.045	.094	.052	.038	.050
2 (N = 90)	.073	.000	.066	.063	.061	.087	.061	.070	.064	.063	.090	.076	.073	.076
3 (N = 205)	.044	.066	.000	.017	.020	.041	.023	.030	.024	.024	.084	.036	.038	.029
4 (N = 151)	.037	.063	.017	.000	.010	.043	.017	.029	.021	.022	.072	.034	.028	.031
5 (N = 165)	.037	.061	.020	.010	.000	.040	.016	.024	.015	.019	.079	.031	.021	.017
6 (N = 137)	.066	.087	.041	.043	.040	.000	.046	.051	.045	.044	.107	.057	.058	.057
7 (N = 159)	.044	.061	.023	.017	.016	.046	.000	.027	.020	.021	.083	.033	.035	.031
8 (N = 172)	.044	.070	.030	.029	.024	.051	.027	.000	.024	.026	.089	.016	.042	.033
9 (N = 112)	.044	.064	.024	.021	.015	.045	.020	.024	.000	.017	.085	.031	.037	.017
10 (N = 158)	.045	.063	.024	.022	.019	.044	.021	.026	.017	.000	.084	.032	.035	.030
11 (N = 103)	.094	.090	.084	.072	.079	.107	.083	.089	.085	.084	.000	.095	.071	.094
12 (N = 82)	.052	.076	.036	.034	.031	.057	.033	.016	.031	.032	.095	.000	.038	.043
13 (N = 88)	.038	.073	.038	.028	.021	.058	.035	.042	.037	.035	.071	.038	.000	.049
14 (N = 102)	.050	.076	.029	.031	.017	.057	.031	.033	.017	.030	.094	.043	.049	.000

Table A-III

Distance (d^2) Matrix of Sevier County, Tennessee: 1880

Districts	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1 (N = 284)	.000	.052	.048	.039	.038	.043	.049	.047	.044	.044	.092	.053	.040	.047	.052
2 (N = 207)	.052	.000	.043	.033	.030	.035	.041	.044	.035	.035	.062	.036	.035	.041	.045
3 (N = 269)	.048	.043	.000	.017	.023	.022	.029	.030	.023	.023	.080	.033	.036	.025	.023
4 (N = 207)	.039	.033	.017	.000	.009	.018	.014	.023	.014	.016	.060	.026	.020	.020	.017
5 (N = 323)	.038	.030	.023	.009	.000	.015	.016	.021	.013	.014	.068	.024	.015	.014	.018
6 (N = 248)	.043	.035	.022	.018	.015	.000	.024	.023	.016	.013	.071	.025	.028	.020	.026
7 (N = 230)	.049	.041	.029	.014	.016	.024	.000	.025	.018	.022	.077	.028	.034	.020	.008
8 (N = 243)	.047	.044	.030	.023	.021	.023	.025	.000	.019	.023	.080	.019	.034	.019	.029
9 (N = 223)	.044	.035	.023	.014	.013	.016	.018	.019	.000	.012	.072	.021	.027	.011	.023
10 (N = 282)	.044	.035	.023	.016	.014	.013	.022	.023	.012	.000	.071	.025	.028	.017	.026
11 (N = 211)	.092	.062	.080	.060	.068	.071	.077	.080	.072	.071	.000	.078	.052	.077	.078
12 (N = 158)	.053	.036	.033	.026	.024	.025	.028	.019	.021	.025	.078	.000	.036	.026	.032
13 (N = 148)	.040	.035	.036	.020	.015	.028	.034	.034	.027	.028	.052	.036	.000	.031	.037
14 (N = 190)	.047	.041	.025	.020	.014	.020	.020	.019	.011	.017	.077	.026	.031	.000	.023
15 (N = 123)	.052	.045	.023	.017	.018	.026	.008	.029	.023	.026	.078	.032	.037	.023	.000

VITA

Joseph Carl Lewelling, a native of Sevier County, Tennessee, was born on January 1, 1959 of a non-consanguineous marriage. In December of 1983 he received a Bachelor of Arts in Liberal Arts with a major in Anthropology from the University of Tennessee, Knoxville. In 1985 he began study toward a Master of Arts in Anthropology at the University of Tennessee.